



# **Hierarchical Image Classification over Visual Tree**

---

**Jianping Fan**

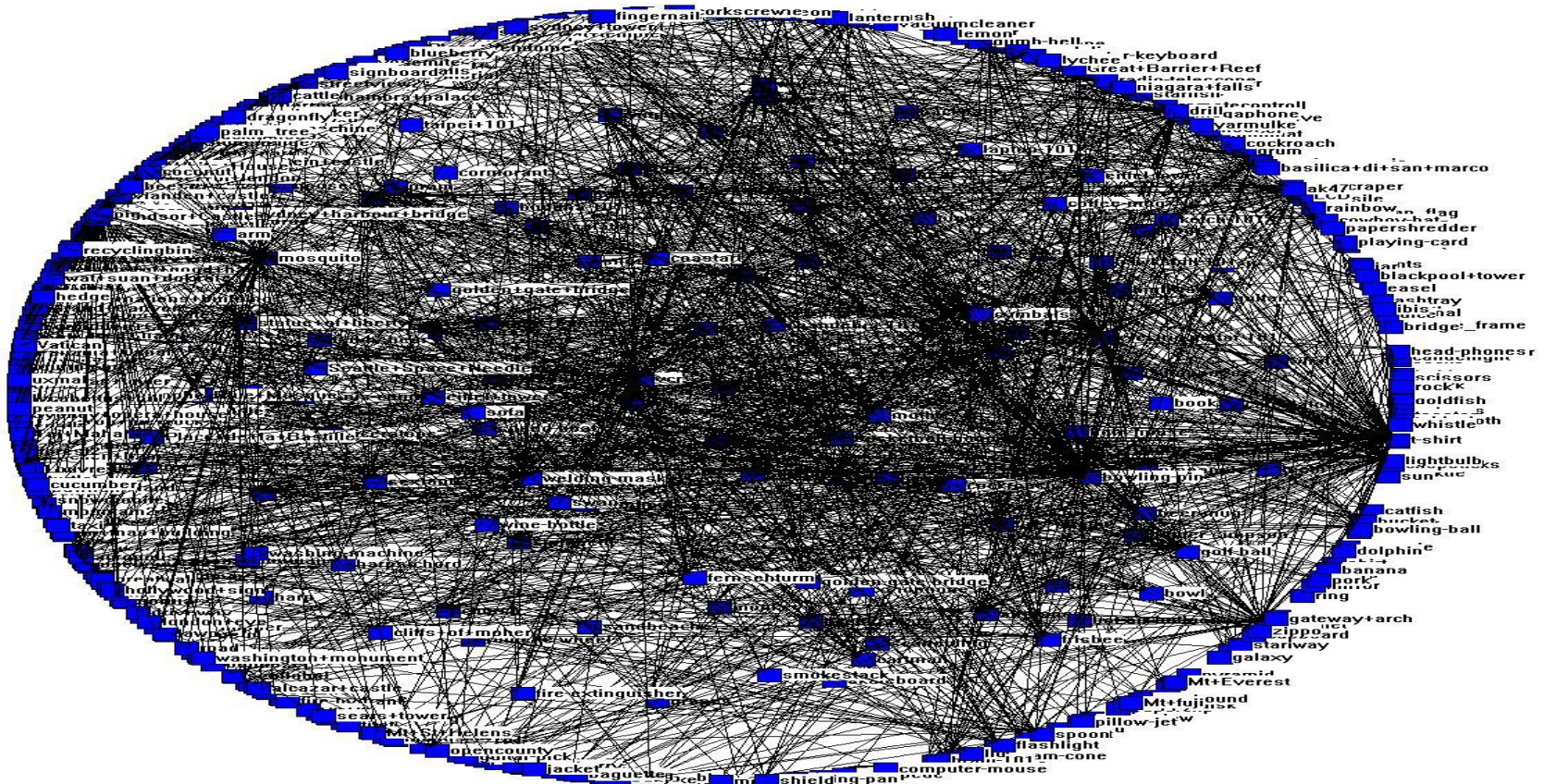
**Dept of Computer Science**

**UNC-Charlotte, NC 28223**

**[www.cs.uncc.edu/~jfan](http://www.cs.uncc.edu/~jfan)**

# 1. Problems to be addressed

## Large-Scale Visual Recognition



**Inter-Object Visual Correlations rather than independency**



## 1. Problems to be addressed

---

- **Large-Scale Visual Recognition: Challenges**
  - **We need to learn large amounts of classifiers for large-scale visual recognition!**
  - **Some object classes and image concepts are visually-related and hard to be discriminated!**
  - **Some object classes and image concepts may have huge inner-concept visual diversity!**



## 1. Problems to be addressed

---

- **Large-Scale Visual Recognition: Challenges**
  - **Huge inner-concept visual diversity**
    - simple models may not work, but using complex models may overlap with others!
  - **Huge inter-concept visual similarity**
    - training complexity will increase for distinguishing visually-related concepts!
  - **Huge computational cost**
    - thousands of inter-related classifiers should be trained jointly!



## 1. Problems to be addressed

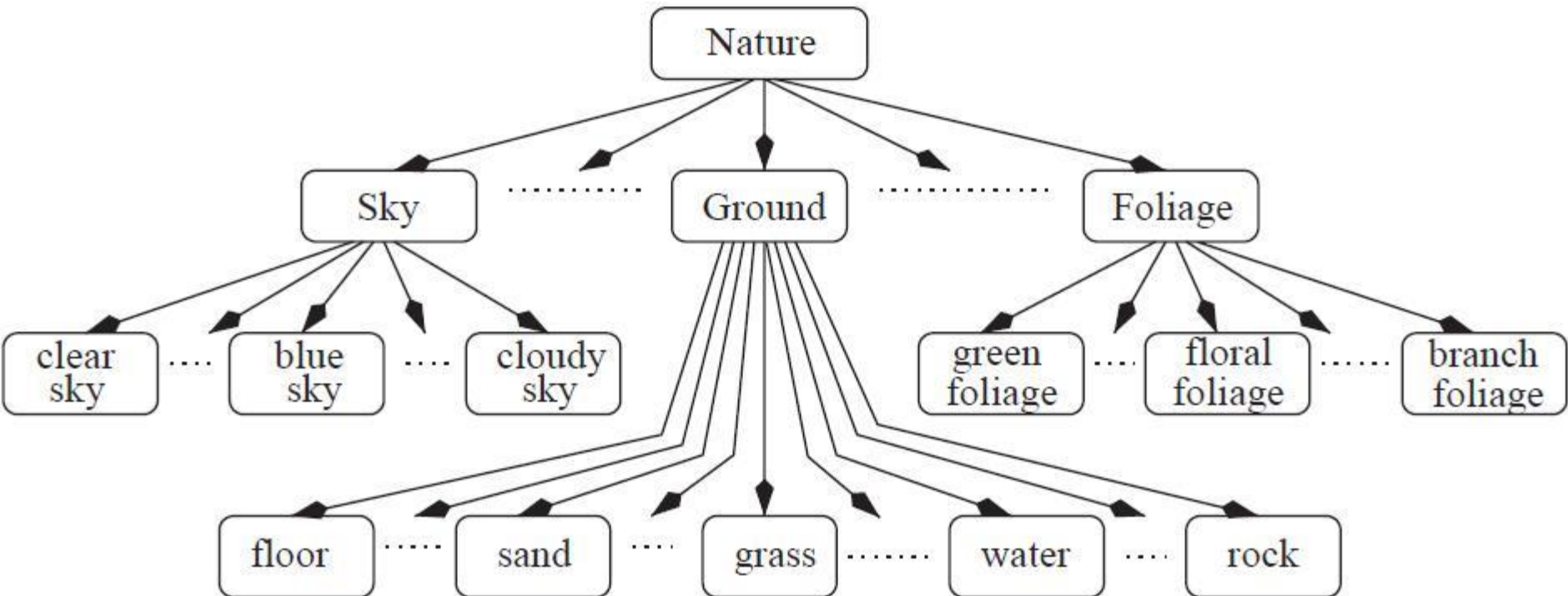
---

### ■ **Large-Scale Visual Recognition**

- **How to leverage social images for classifier training?**
- **How to leverage inter-class correlations for dictionary learning & classifier training?**
- **How to deal with inter-level error propagation?**

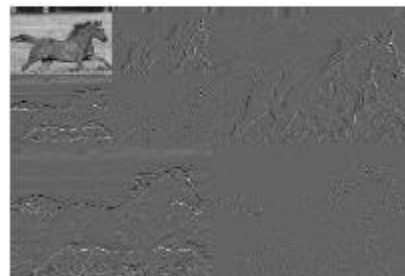
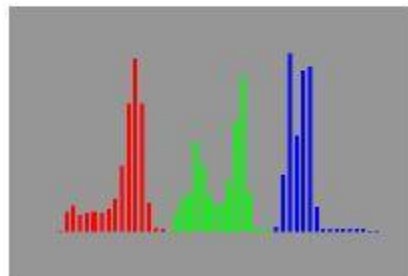
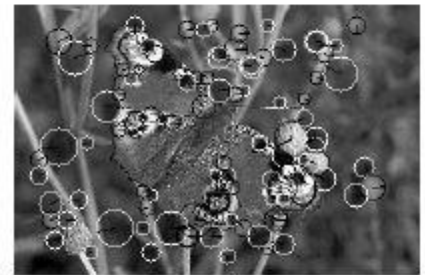
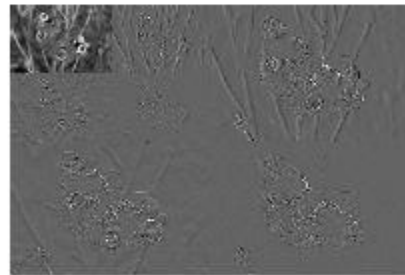
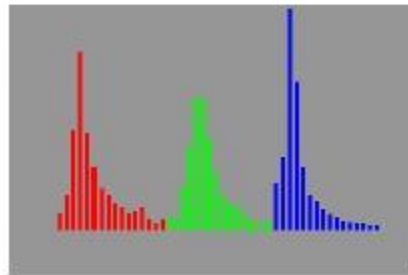
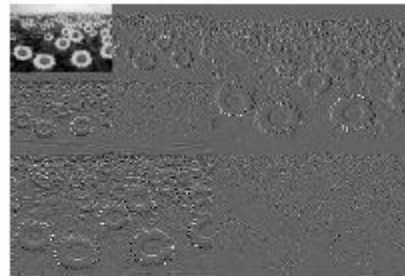
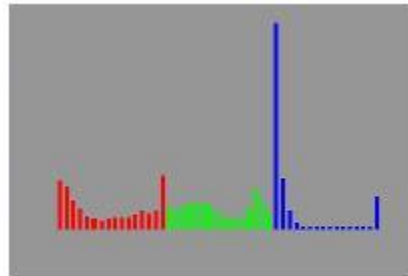
## 2. Collecting Large-Scale Training Images

- **Flickr Images & Other Image Sites**
  - **Keywords for image crawling**



## 2. Collecting Large-Scale Training Images

### ■ Visual Feature Extraction



## 2. Collecting Large-Scale Training Images

- **Synonymous Concepts: Visual Similarity**



(a) Auto



(b) Automobile



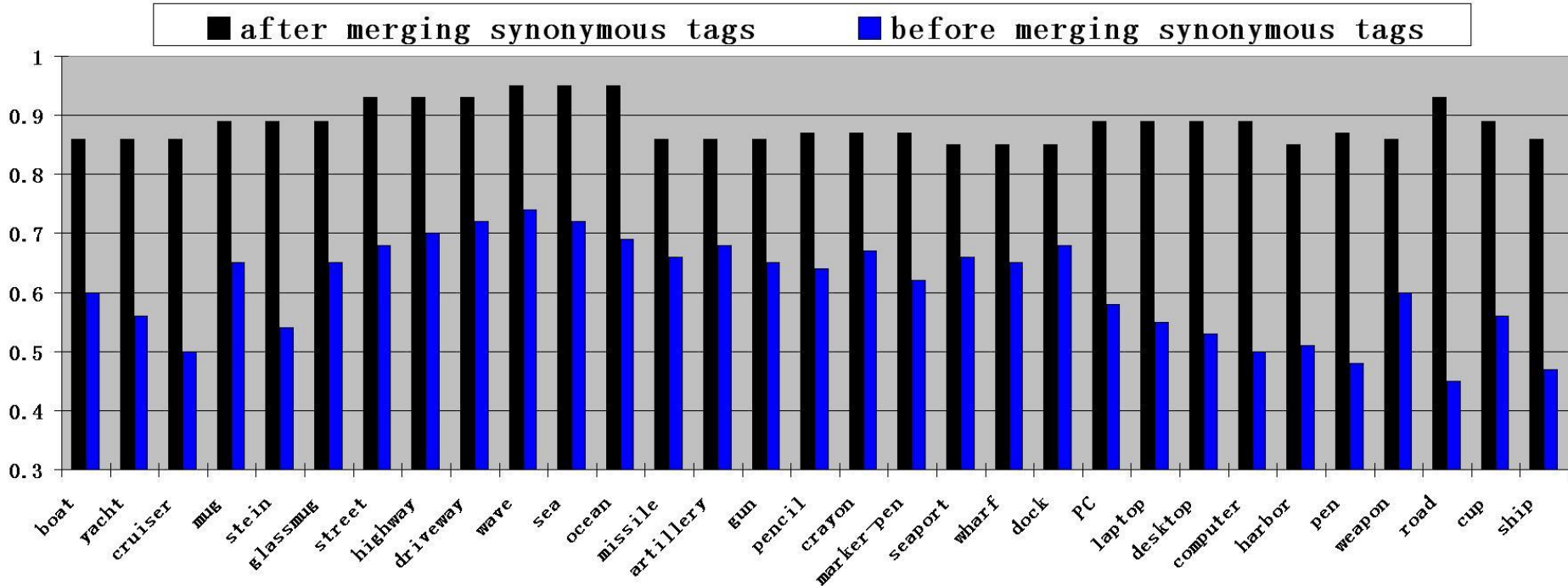
(c) Car

CVPR2010



## 2. Collecting Large-Scale Training Images

### ■ Synonymous Concepts: Visual Similarity

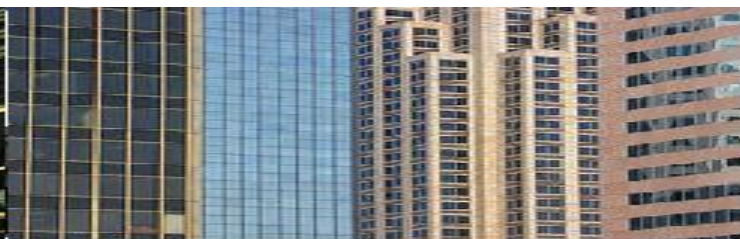


## 2. Collecting Large-Scale Training Images

- **Ambiguous Concept: Visual Diversity**



**(a) Bank Office**



**(b) River Bank**

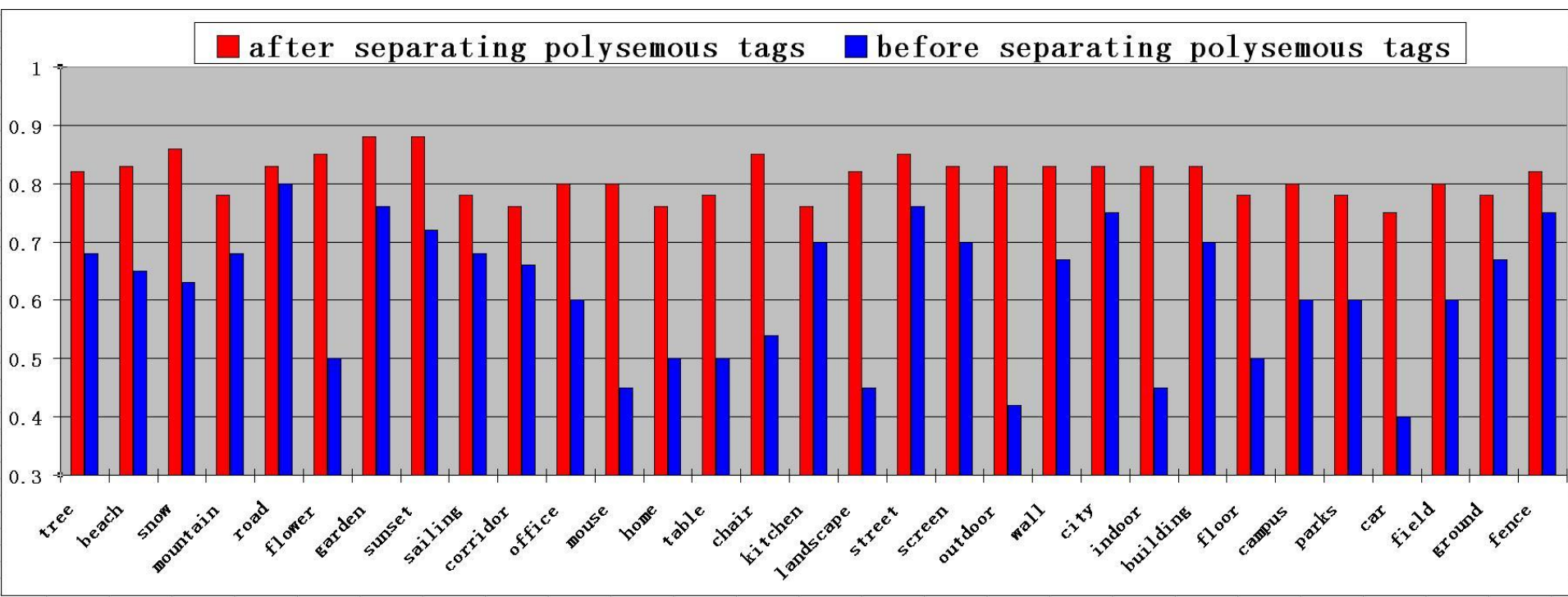


**(c) Cloud Bank**



## 2. Collecting Large-Scale Training Images

### ■ Ambiguous Concept: Visual Diversity



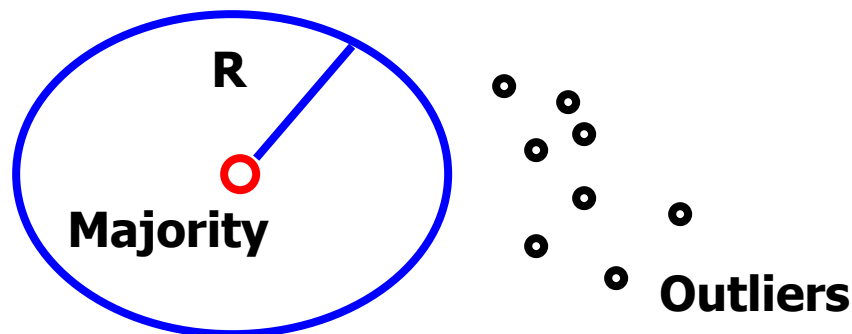
## 2. Collecting Large-Scale Training Images

### ■ Junk Image Filtering

IEEE Trans. CSVT 2009

$$\forall_{j=1}^N : \|\phi(x_j) - \mu^\phi\|^2 \leq R^2$$

$$\mu^\phi = \sum_{j=1}^N \phi(x_j)$$



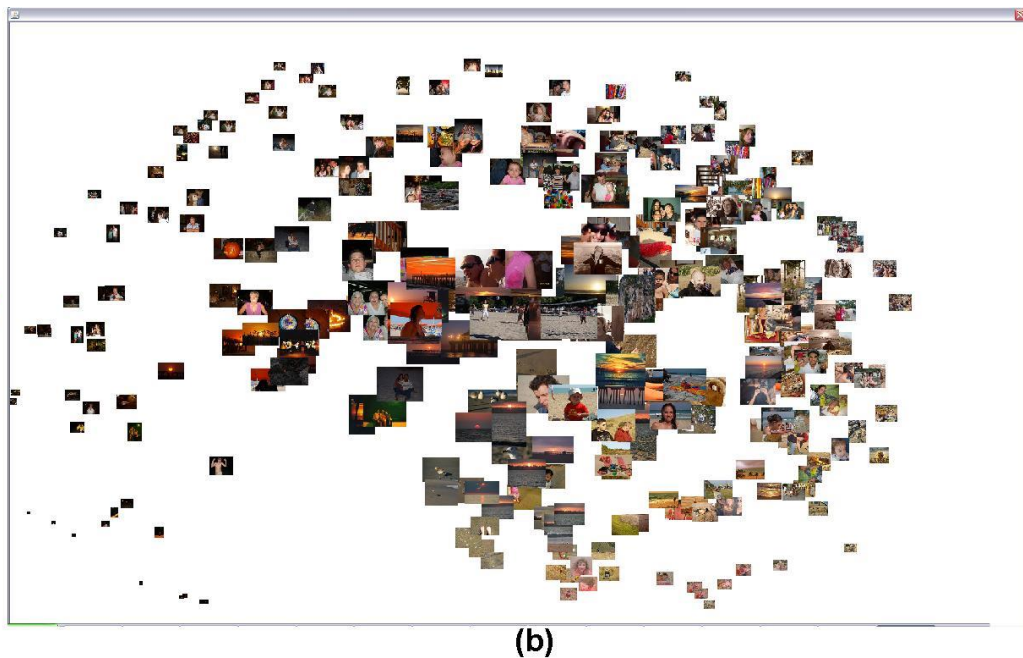
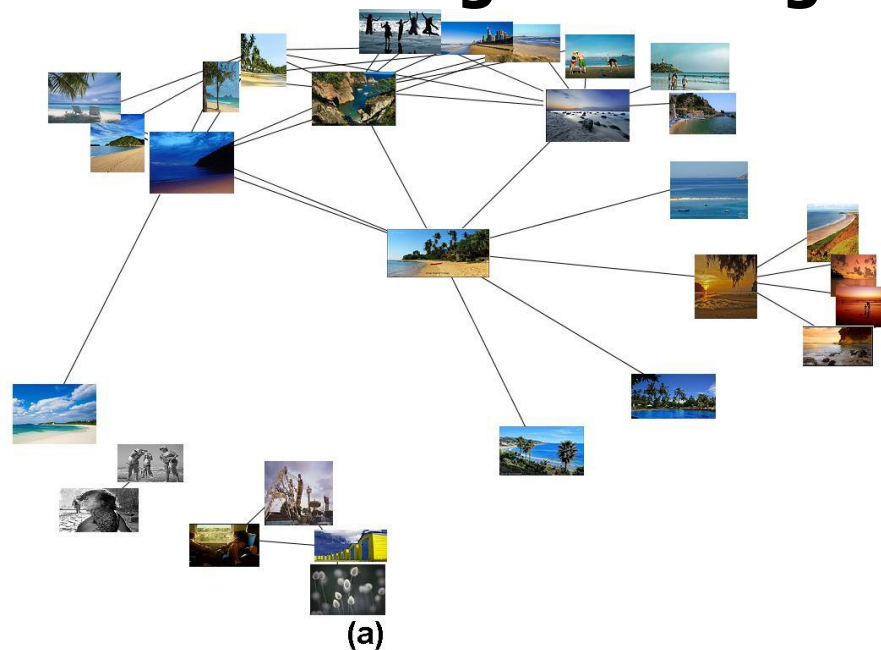
$$\min \left\{ R^2 + \frac{C}{N} \sum_{j=1}^N \xi_j \right\} \quad \text{subject to:} \quad \forall_{j=1}^N : \|\phi(x_j) - \mu^\phi\|^2 \leq R^2 + \xi_j, \quad \xi_j \geq 0$$

**Decision function:**

$$f(x) = R^2 - \sum_{i,j} \alpha_i \alpha_j \kappa(x_i, x_j) + 2 \sum_j \alpha_j \kappa(x_j, x) - \kappa(x, x)$$

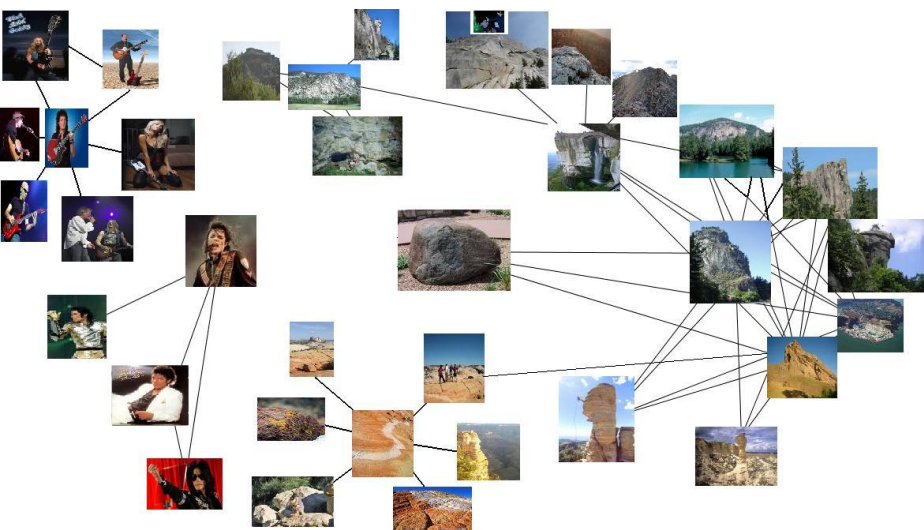
## 2. Collecting Large-Scale Training Images

### ■ Junk Image Filtering

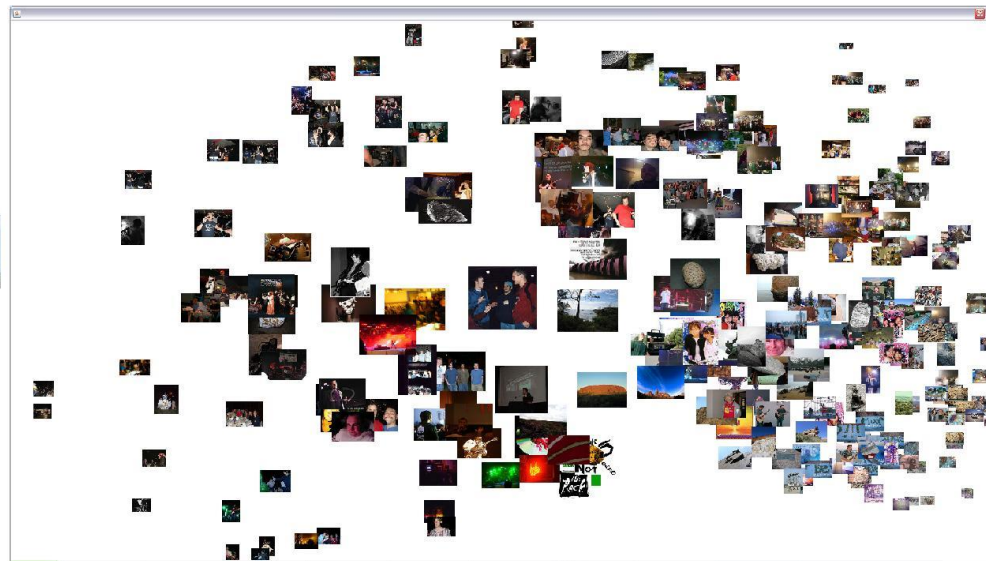


## 2. Collecting Large-Scale Training Images

### ■ Junk Image Filtering



(a)



(b)

Most text terms are weakly related or even irrelevant to web images in the same webpage

Research Pictures Videos Zoos Forum

### Tiger Pictures, Information and Facts

The majestic Tiger is the largest of the big cats. Today Tigers are only found in South and Southeast Asia, China and the Russian Far East. Once there were nine subspecies of tigers: Bengal, Siberian, Indochinese, South Chinese, Sumatran, Malayan, Caspian, Javan and Bali. The last three of these are extinct, one is extinct in the wild, and the rest are on the endangered species list. In the Wild Tigers have a lifespan of 10 to 15 years. In Captivity they have been known to live up to 20 years.



Q&A

[Ads by Google](#) [Black Tiger](#) [Tiger Woods](#) [Tiger](#) [Tiger to Lion](#)

[White Tiger Facts](#)  
[Where do Tigers Live?](#)  
[Mating Habits of Tigers](#)  
[Extinct Tigers](#)  
[How Big is a Tiger?](#)  
[What is a Caspian Tiger?](#)

Most Tigers have long, thick reddish coats with white bellies and white and black tails with narrow stripes of black, brown or gray on their heads, bodies, tails and limbs. They have round pupils and yellow irises with the exception of the white tigers who generally have blue eyes. They have powerful jaws, shoulders and legs. Tigers range in length from 6.6 feet 2 feet of which is tail to 12.2 feet 3 feet of which is tail. They weigh up to 700 pounds for the largest Siberian Tigers to only 200 pounds for the smallest Sumatran Tigers.

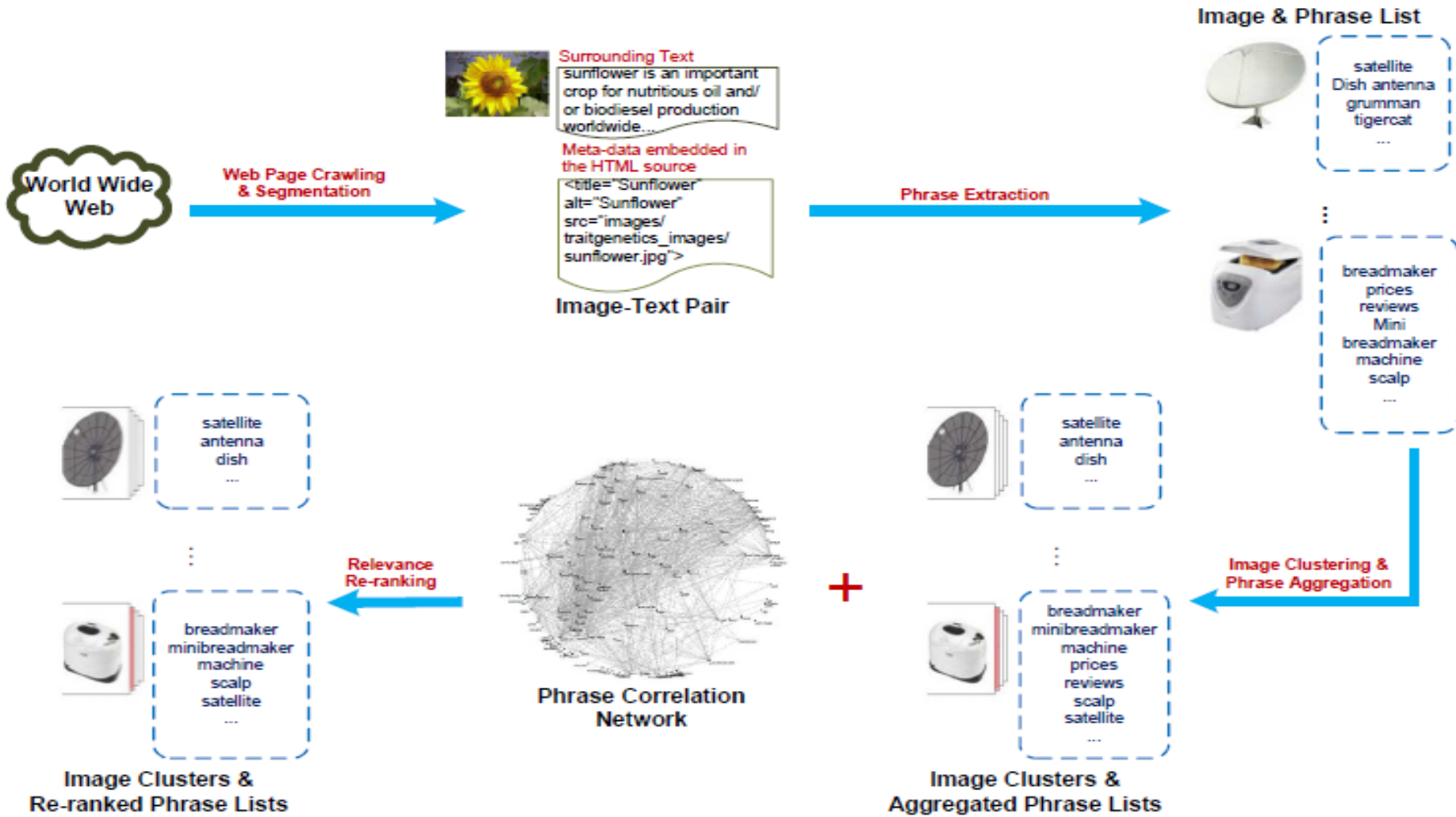


The Tiger's stripes are a very useful camouflage, hiding the T in long grass or reeds. Tigers are mostly nocturnal, active at night, and rely on that camouflage to ambush their prey. Tige are the most powerful of the **big cats**. They can take down larger than themselves. They use their massive body weight t knock prey to the ground and kill with powerful jaws to the neck. Tigers are excellent swimmers and have been known to carry 300 pound animals with them while swimming. Dependin their environment Tigers eat monkeys, hares, deer, wild pigs, water buffalo, antelope, sloth

Noise image

tiger  
big cat  
Southeast Asia  
Russian  
Chinese  
Bengal  
Siberian  
Indochinese  
South Chinese  
.....

# Text-Image Alignment for Web Image Indexing





# Informative Image Extraction

The screenshot shows the HowStuffWorks website interface. At the top, there is a navigation bar with categories like Adventure, Auto, Culture, Entertainment, Home & Garden, Money, Science, Tech, Video, Shows, Blogs, Quizzes, Games, and Random Article. Below this is a search bar and a secondary navigation bar with categories like Engineering, Environmental Science, Forces of Nature, Innovation, Life Science, Military, Physical Science, Dictionary, Science Vs. Myth, Space, Transportation, and Biology. The main content area features several article thumbnails: 'MORE STUFF LIKE THIS' with a person's face, 'Stuff of Genius: 10 Accidental Inventions You Won't Believe' with eggs, '10 Complete Falsehoods About Food' with a bowl of food, and 'Science Puzzles: Aurora Borealis' with a green aurora. Below these is a section titled 'Classic Airplanes' by the Editors of Publications International, Ltd. This section includes a large image of a biplane (highlighted with a red box) and a text block. To the right of the biplane image is a 'Flight Image Gallery' section with a list of items: '1 10 Absolutely Worst Foods to Eat', '2 Are you a werewolf?', and '3 10 Grossest Things in Your Body Right Now'. There is also a 'SEE HOW REVERSE LOGISTICS BOOSTS PRODUCT VALUE WHILE KEEPING CUSTOMERS HAPPY.' advertisement with a 'VIEW REPORT' button. A 'Noise image' label with arrows points to several elements: the 'Aurora Borealis' image, the '10 Complete Falsehoods About Food' image, the 'VIEW REPORT' button, and the '10 Grossest Things in Your Body Right Now' image.

Noise image

- Two simple rules:
  - Aspect ratio ( $>0.2$  or  $<5$ )
  - Image size ( $\min(\text{width}, \text{height}) > 60$  pixel)
- Not perfect but produce satisfied results
- Unsupervised and computationally efficient

# Webpage Segmentation

## ■ Surrounding Text Extraction

### Classic Airplanes

by the Editors of Publications International, Ltd.

Share 9 Like 12 Tweet 2 +1 0

Page 1 2 3 4 ▶



French fighter pilots who were seriously challenged by top-ranked German airplanes during World War I welcomed the rugged SPAD VII. See more flight pictures.

In a sense, all airplanes are classic airplanes, because each one represents the very best its designer and builder could do, given the talent, materials, and time available at the moment. No development group ever set out to make a second-best airplane. Instead, every aircraft, and especially every classic aircraft described and pictured within these pages, was the product of the loving care of an intelligent design team.

#### Flight Image Gallery

The following pages in this article provide links to profiles of classic

airplanes built over the last century. You'll begin with classic airplanes of the Early Years, including the Wright Flyers first successful flight. Learn about the military fighter airplanes of World War I and World War II, and explore the aircraft built during the Golden Age of Flight. Then fast-forward to the present day Jet Age and see how much classic airplanes have progressed over the last 100 years.

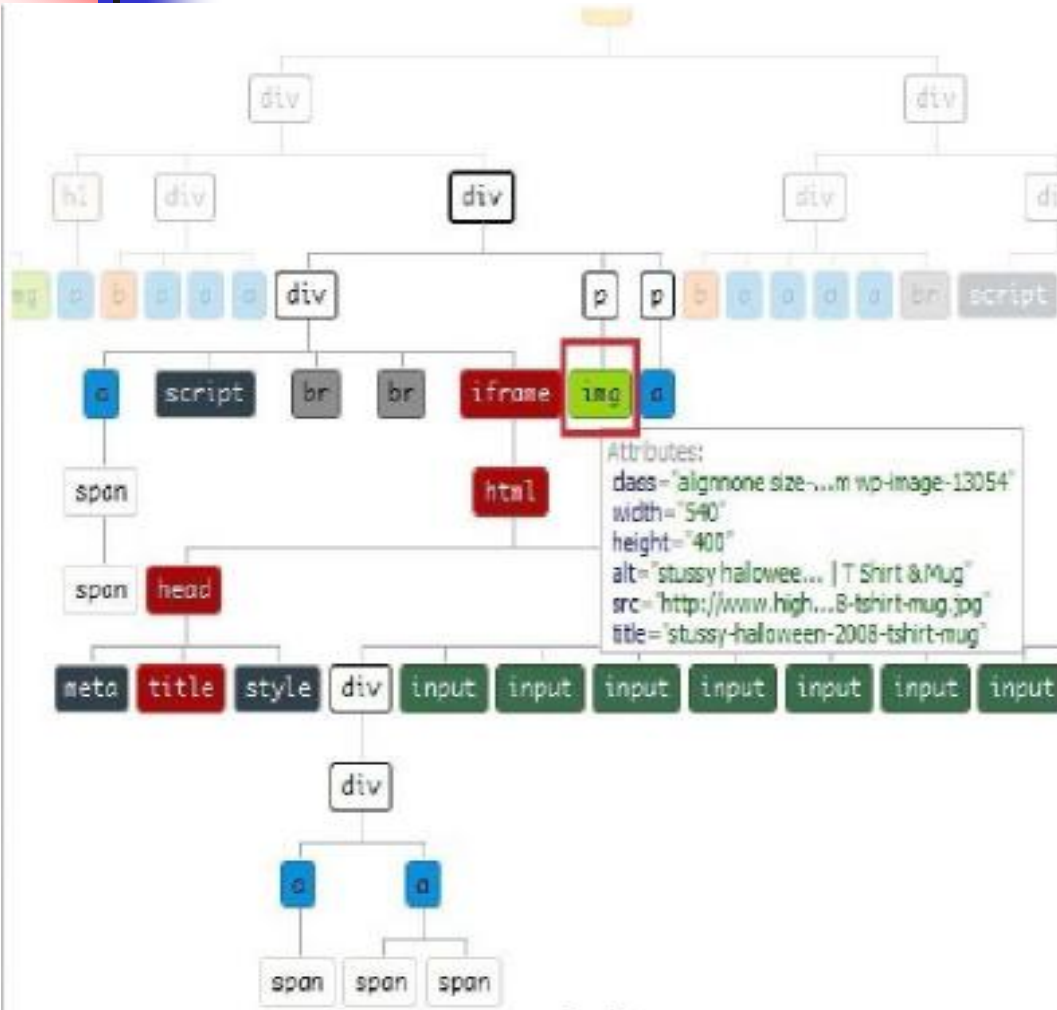


Each and every one of these

- Visual-based algorithm
  - precise but expensive
  - [Cai et al. MSR-TR'03]
- DOM (Document Object Model) based method
  - computationally efficient

# Webpage Segmentation

## Surrounding Text Extraction



- Visual-based algorithm
  - precise but expensive
  - [Cai et al. MSR-TR'03]
- DOM (Document Object Model) based method
  - computationally efficient

# Text-Image Alignment for Web Image Indexing

Cluster No.: 3598, 10 duplicates



**Phrase list 1:** sterilization equipment, water, sterilizer, china mainland  
**Phrase list 2:** autoclave, sterilizer, water, china mainland, manufacturer  
**Phrase list 3:** retort, heating, sterilizer, water, china mainland, manufacturer  
**Phrase list 4:** sterilizer, water, china mainland, manufacturer  
**Phrase list 5:** sterilization equipment, water, sterilizer, china mainland, manufacturer  
....  
**Aggregation:** sterilizer, sterilization equipment, water, retort, manufacturer, ....

Cluster No.: 6244, 13 duplicates



**Phrase list 1:** cimarron, roper, saddle, roper saddle, horse, ...  
**Phrase list 2:** cimarron, roper, saddle, roper saddle,...  
**Phrase list 3:** saddle, roper, roper saddle, horse, sale  
**Phrase list 4:** roper saddle, saddle, cimarron, horse  
....  
**Aggregation:** saddle, roper, roper saddle, cimarron, ....

Cluster No.: 16263, 33 duplicates



**Phrase list 1:** face, area, drive stick, rule safety  
**Phrase list 2:** face, grip, play tennis, tennis racket  
**Phrase list 3:** face, , tennis racket, maintenance  
**Phrase list 4:** face, shaver, tennis preparation tip.,  
....  
**Aggregation:** face, shaver, gillete, ....

Cluster No.: 29906, 8 duplicates



**Phrase list 1:** pisa feb, pisa, leaning tower, location, photo  
**Phrase list 2:** pisa, leaning tower, location, photo  
**Phrase list 3:** pisa, location, leaning tower, photo  
**Phrase list 4:** pisa, leaning tower, photo....  
**Aggregation:** pisa, learning tower, pisa feb, location, ....

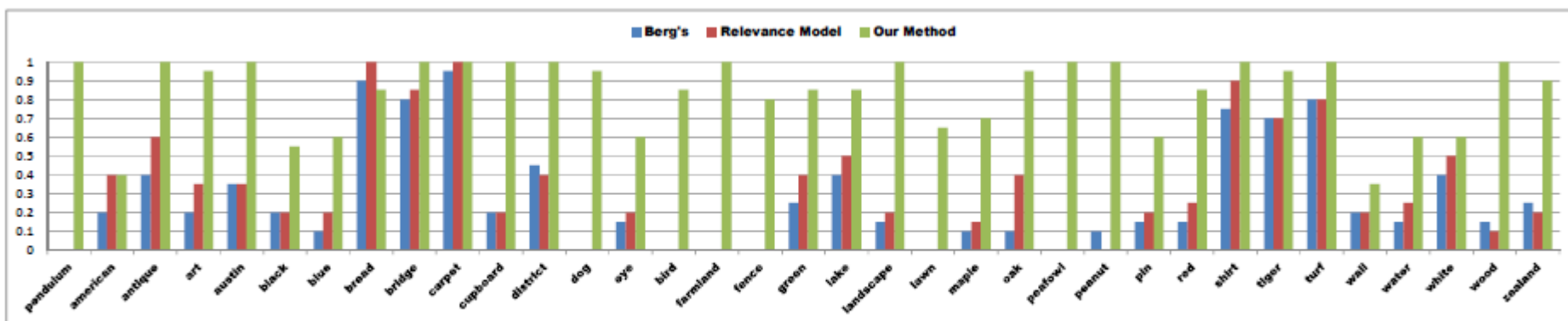
Cluster No.: 35950, 27 duplicates



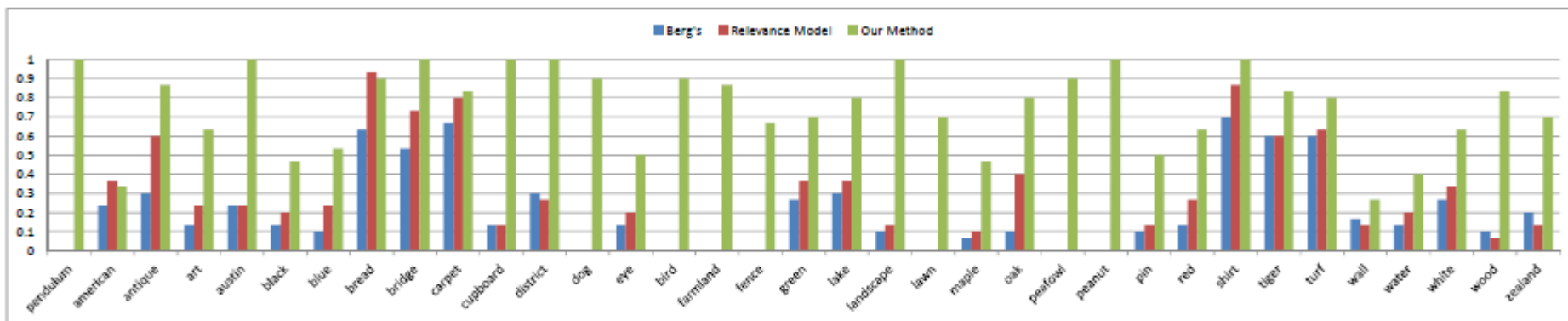
**Phrase list 1:** venture snowmobile, indonesia  
**Phrase list 2:** venture snowmobile, arctic, snowmobile, ...  
**Phrase list 3:** venture snowmobile, snowmobile  
**Phrase list 4:** venture snowmobile, snowmobile manufacture  
**Aggregation:** venture snowmobile, snowmobile, ....

## Near-duplicates share similar semantics!

# Text-Image Alignment for Web Image Indexing



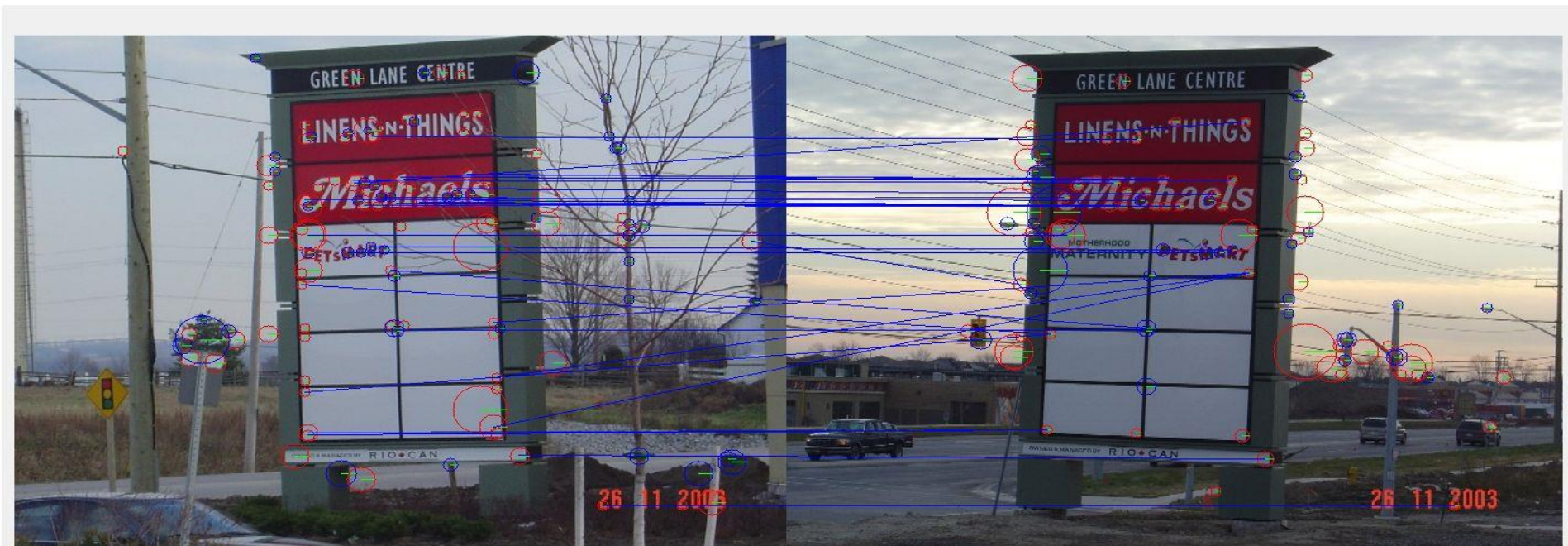
(b) Pre@20, top 20 images are evaluated.



(c) Pre@30, top 30 images are evaluated.

## 2. Collecting Large-Scale Training Images

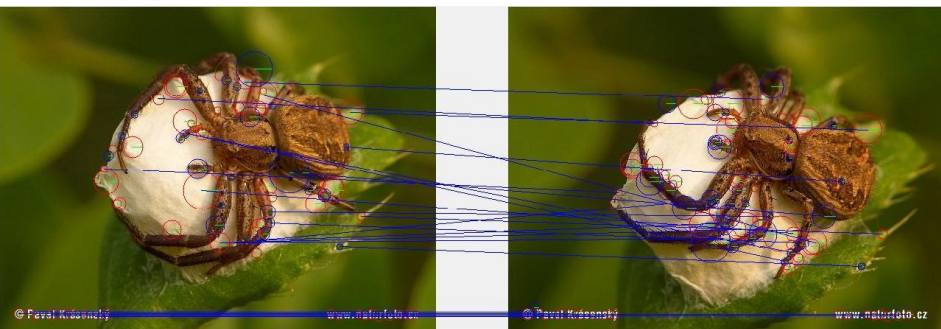
- Duplicate Detection



**Duplicates may mislead classifier training tools!**

## 2. Collecting Large-Scale Training Images

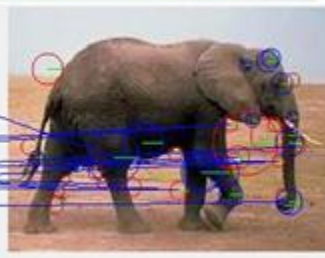
### ■ Duplicate Detection



(a)



(b)



## 2. Collecting Large-Scale Training Images

### Automatic Tag-Instance Alignment



(a) Original Image

Image Tags:

Bush  
Tree  
Grass  
Horse

Tag-Instance  
Alignment



(b) Multiple Image Instances

ACM MM 2010

### Missing Tag Prediction



Image Segmentation &  
Instance-Tag Alignment



Object Co-Occurrence  
Contexts for Missing  
Object Tag Prediction



CVPR 2012





### 3. Visual Concept Network

---

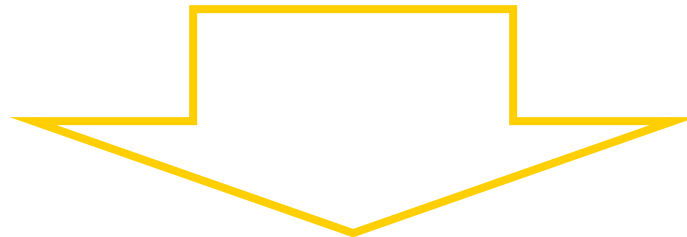
ACM MM2009

- **Why we need visual concept network?**

  - concept ontology, object co-occurrence network, ....

- **Common space: classifier training & concept detection**

  - visual feature space rather than label space or concept space



**We need to characterize inter-concept visual correlations rather than others!**

**Inter-related learning task determination**

### 3. Visual Concept Network

Image Concept: Sunset



Image Concept: Beach



Image Similarity Characterization

Image Similarity Characterization

Inter-Concept Visual  
Similarity Characterization

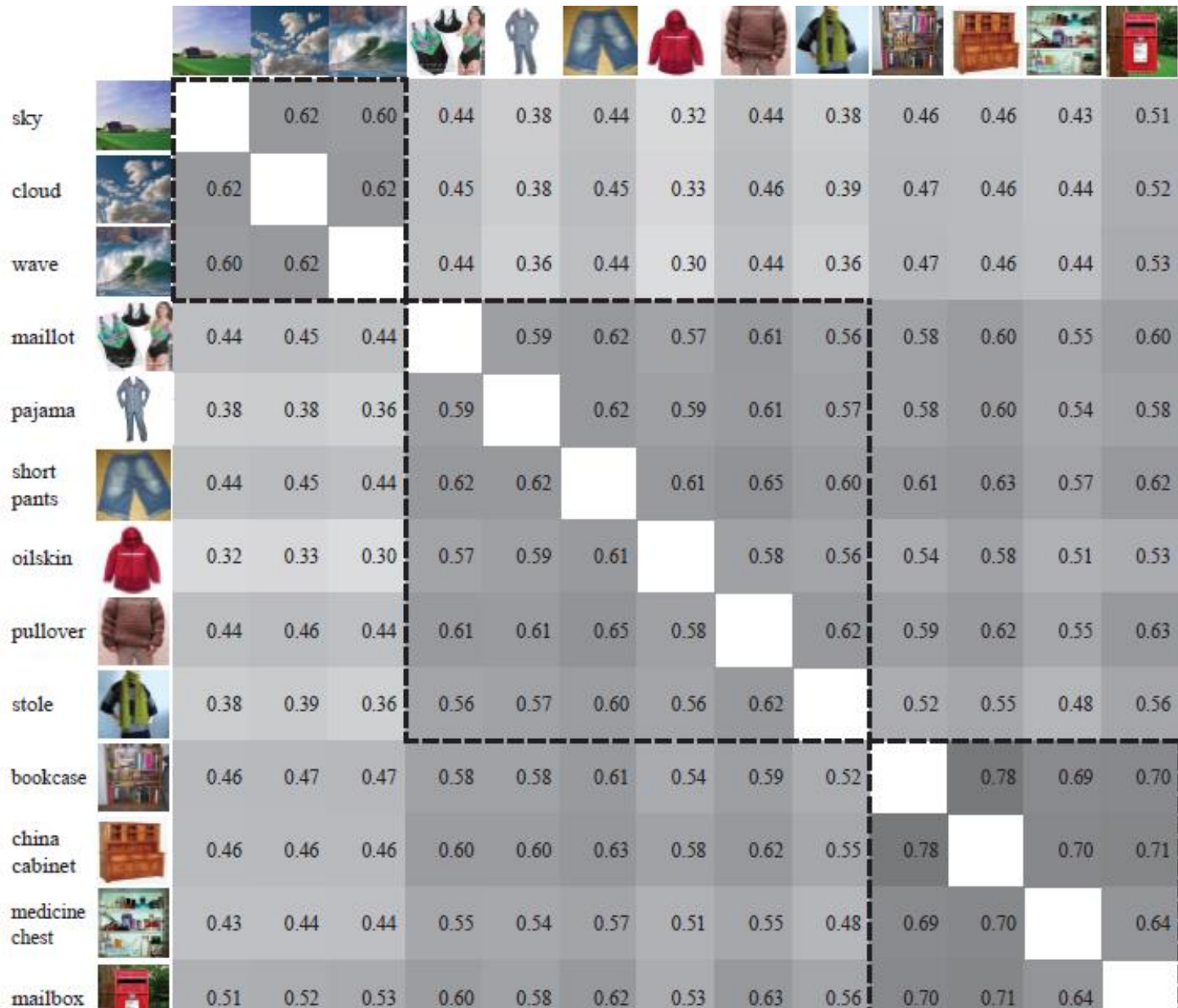


### 3. Visual Concept Network

---

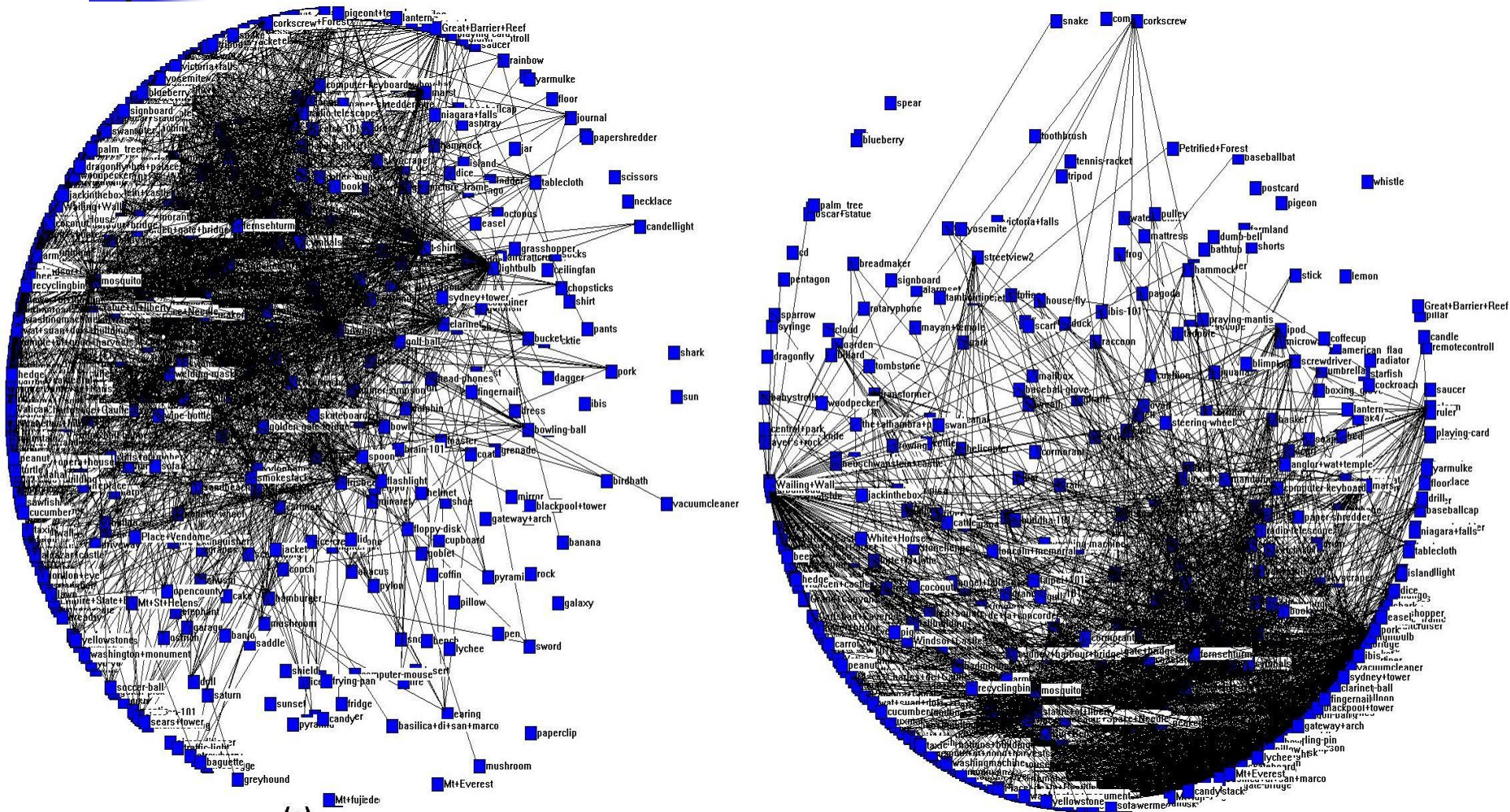
concept pair	$\gamma$	concept pair	$\gamma$	concept pair	$\gamma$	concept pair	$\gamma$
urbanroad-streetview	0.99	cat-dog	0.81	kerb-saucer	0.28	tweezer-corn	0.19
frisbee-pizza	0.80	dolphin-cruiser	0.73	fridge-vest	0.29	journal-grape	0.19
moped-bus	0.75	habor-outview	0.71	stick-cupboard	0.29	sheep-greatwall	0.26
monkey-humanface	0.71	guitar-violin	0.71	mushroom-moon	0.32	whistle-watermelon	0.28
lightbulb-firework	0.69	mango-broccoli	0.69	cannon-ruler	0.41	snake-ipod	0.31
porcupine-lion	0.68	bridge-warship	0.68	tombstone-crab	0.42	helicopter-city	0.63
doorway-street	0.65	statue-building	0.68	pylon-highway	0.61	LCD-container	0.65
windmill-bigben	0.63	cat-lion	0.66	beermug-bar	0.62	sailboat-cruiser	0.66

# 3. Visual Concept Network

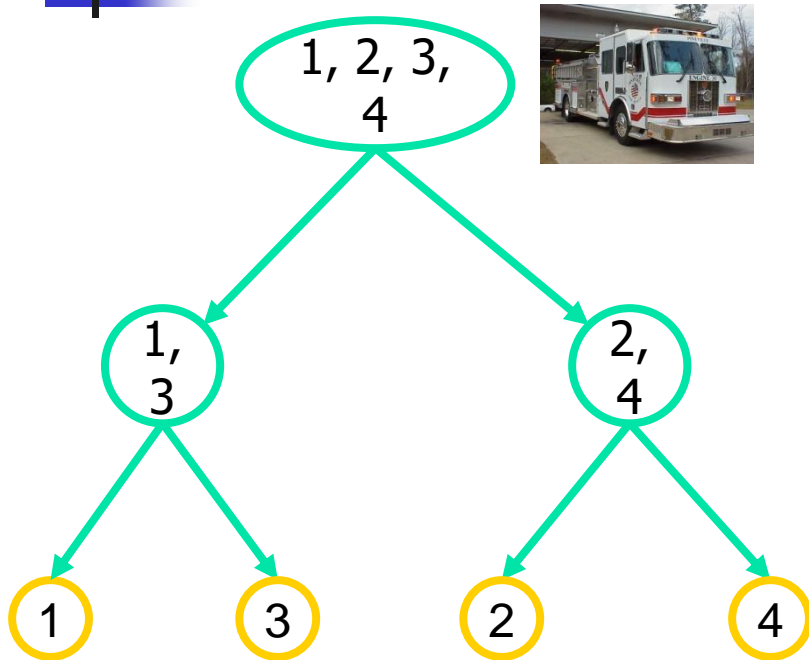




# 3. Visual Concept Network



# Label Tree for Efficient Classification



Label 1: cat  
Label 3: dog

Label 2: mini van  
Label 4: fire truck

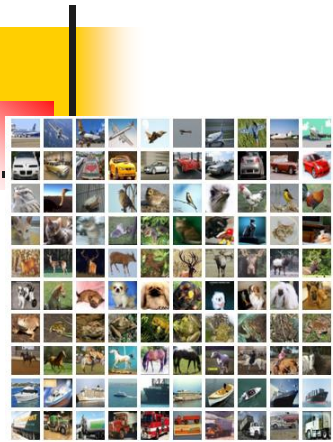
Number of dot products  
needed in the label tree:  
 $1 + 1 = 2$

Number of dot product  
needed in a flat approach:  
 $1 + 1 + 1 + 1 = 4$

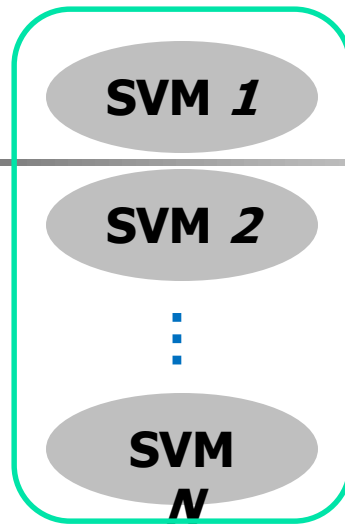
It is a fire truck!

[Bengio et al. [NIPS'2010](#)]

# Construction of Label Tree

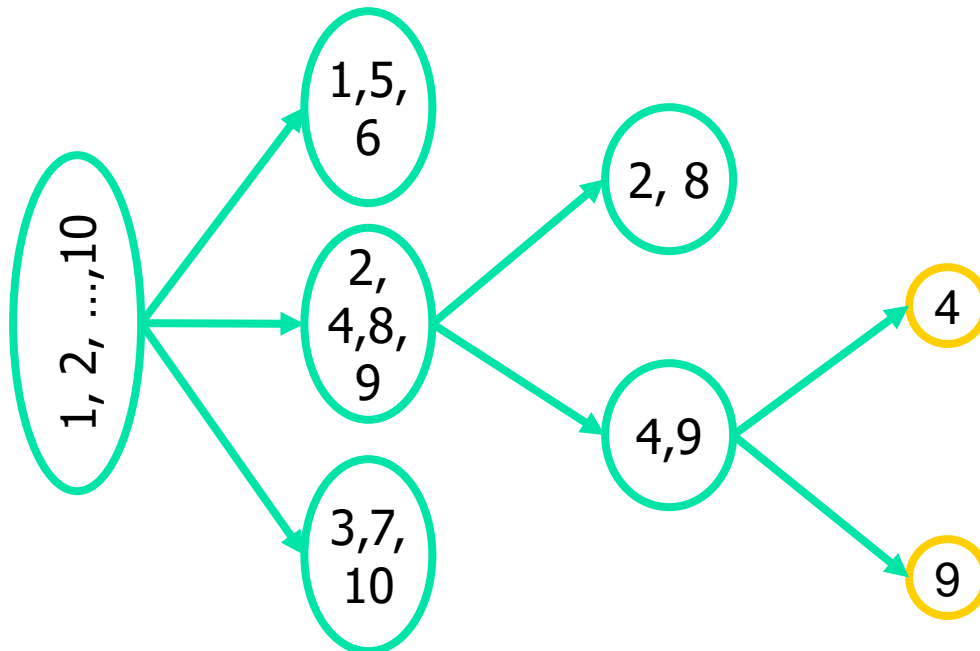


Testing Samples



83.0	0.0	3.0	3.0	1.0	2.0	4.0	0.0	2.0	2.0
0.0	95.0	0.0	0.0	0.0	4.0	0.0	0.0	0.0	1.0
6.0	0.0	71.0	4.0	1.0	2.0	0.0	6.0	2.0	8.0
5.0	1.0	5.0	68.0	5.0	0.0	1.0	5.0	2.0	8.0
3.0	0.0	2.0	4.0	76.0	0.0	3.0	3.0	7.0	2.0
1.0	6.0	1.0	0.0	0.0	92.0	0.0	0.0	0.0	0.0
1.0	0.0	0.0	3.0	1.0	0.0	92.0	1.0	0.0	2.0
0.0	1.0	7.0	9.0	4.0	4.0	0.0	72.0	2.0	1.0
1.0	0.0	3.0	2.0	15.0	1.0	1.0	4.0	70.0	3.0
3.0	1.0	8.0	8.0	2.0	0.0	5.0	1.0	5.0	67.0

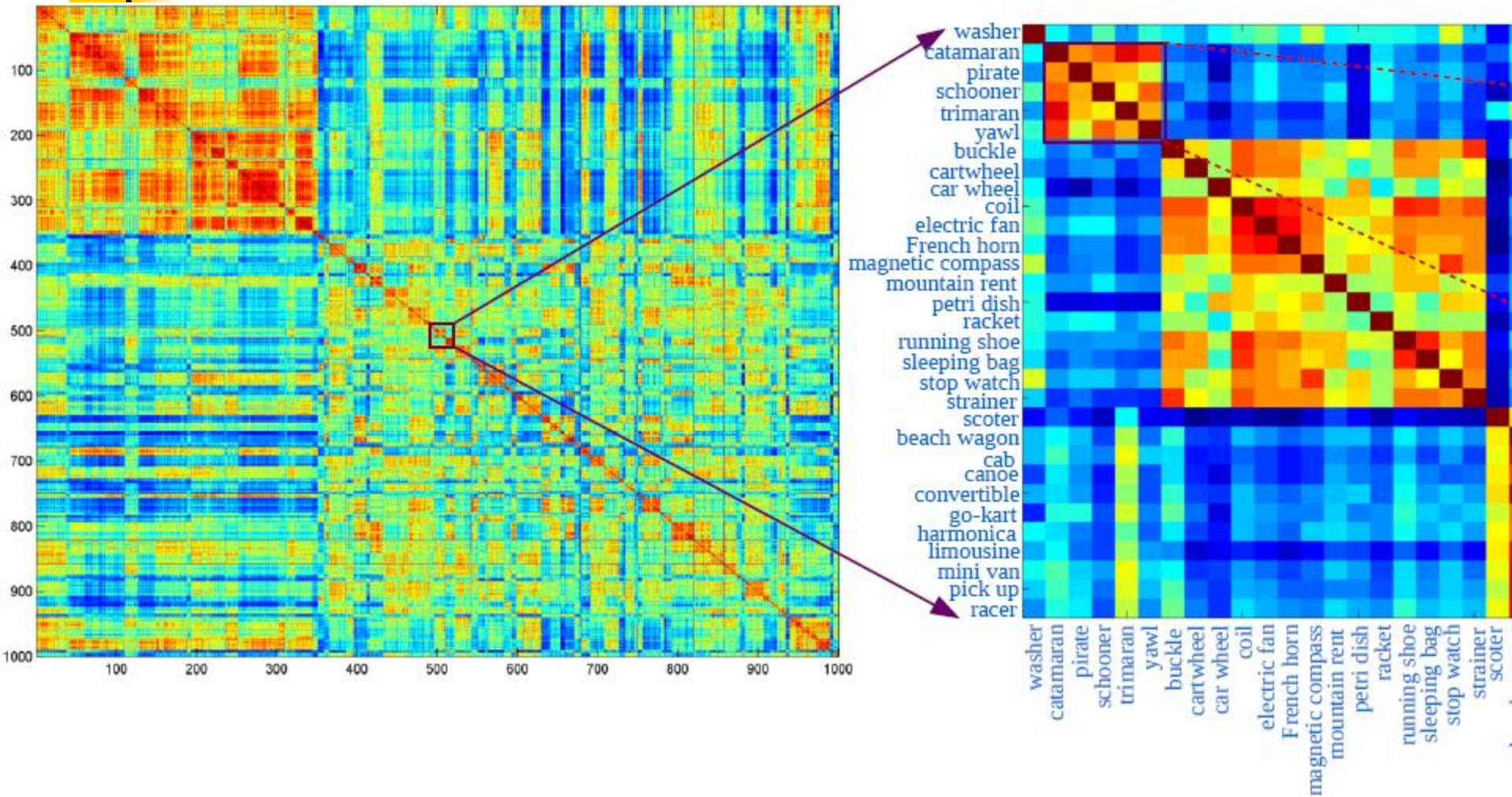
Confusion Matrix



- Training  $N$  one-vs-rest SVMs is very expensive
- The SVMs could be unreliable
  - Huge sample imbalance
  - Negative samples could mislead the classifier training

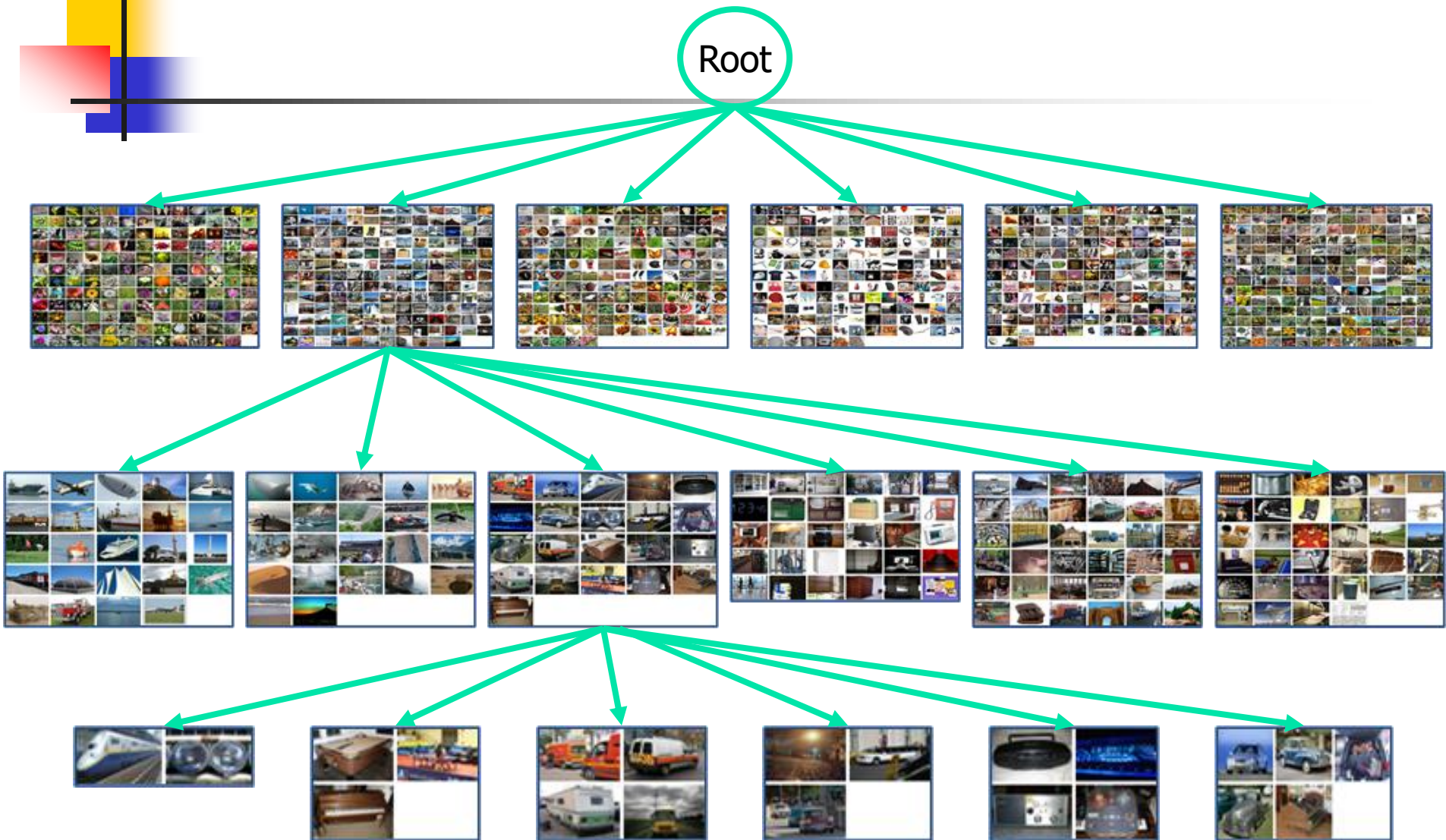


# Visual Similarity Matrix

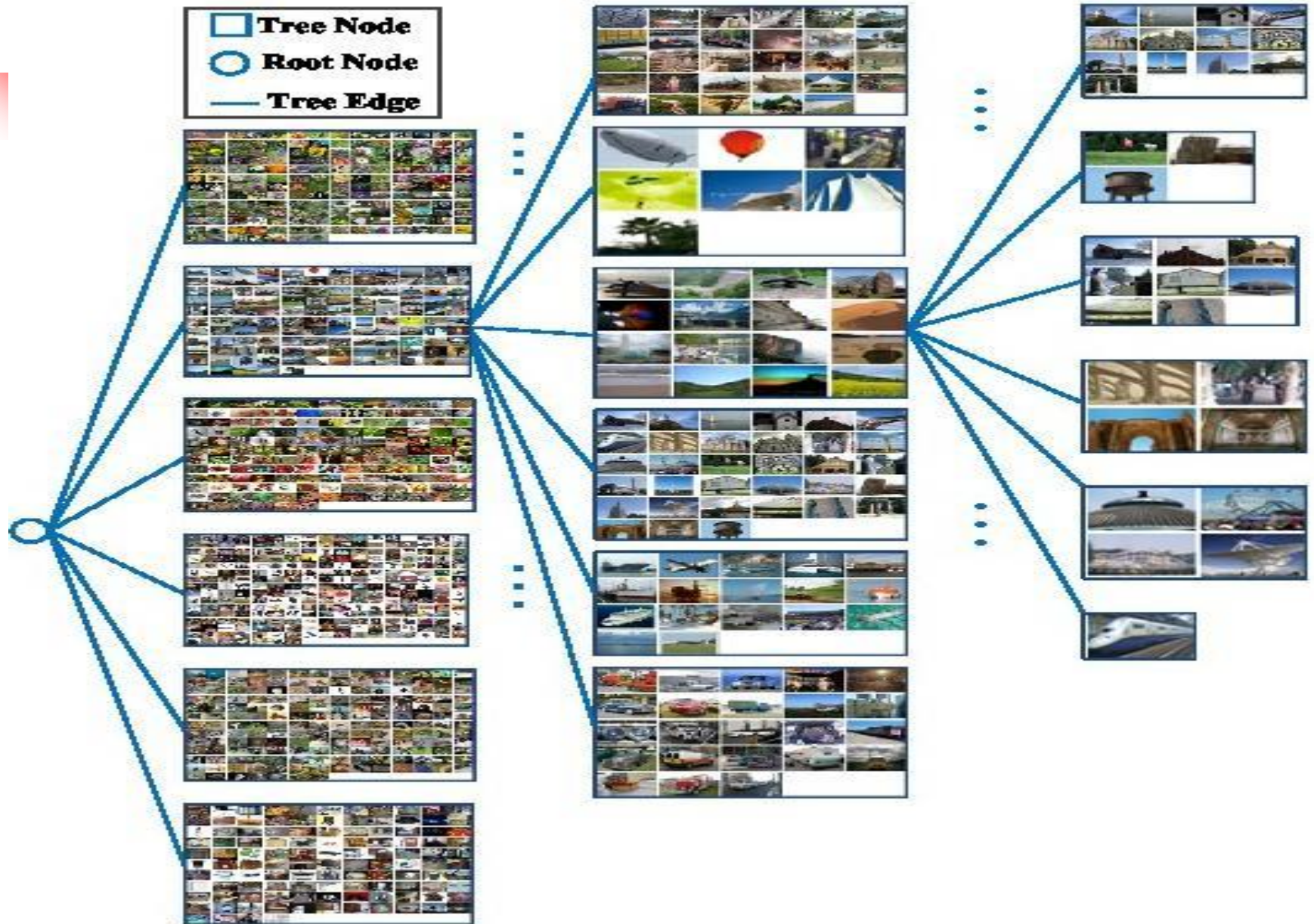


Result is based on ImageNet data set of 1000 categories

# 4. Visual Tree Construction: Hierarchical Clustering



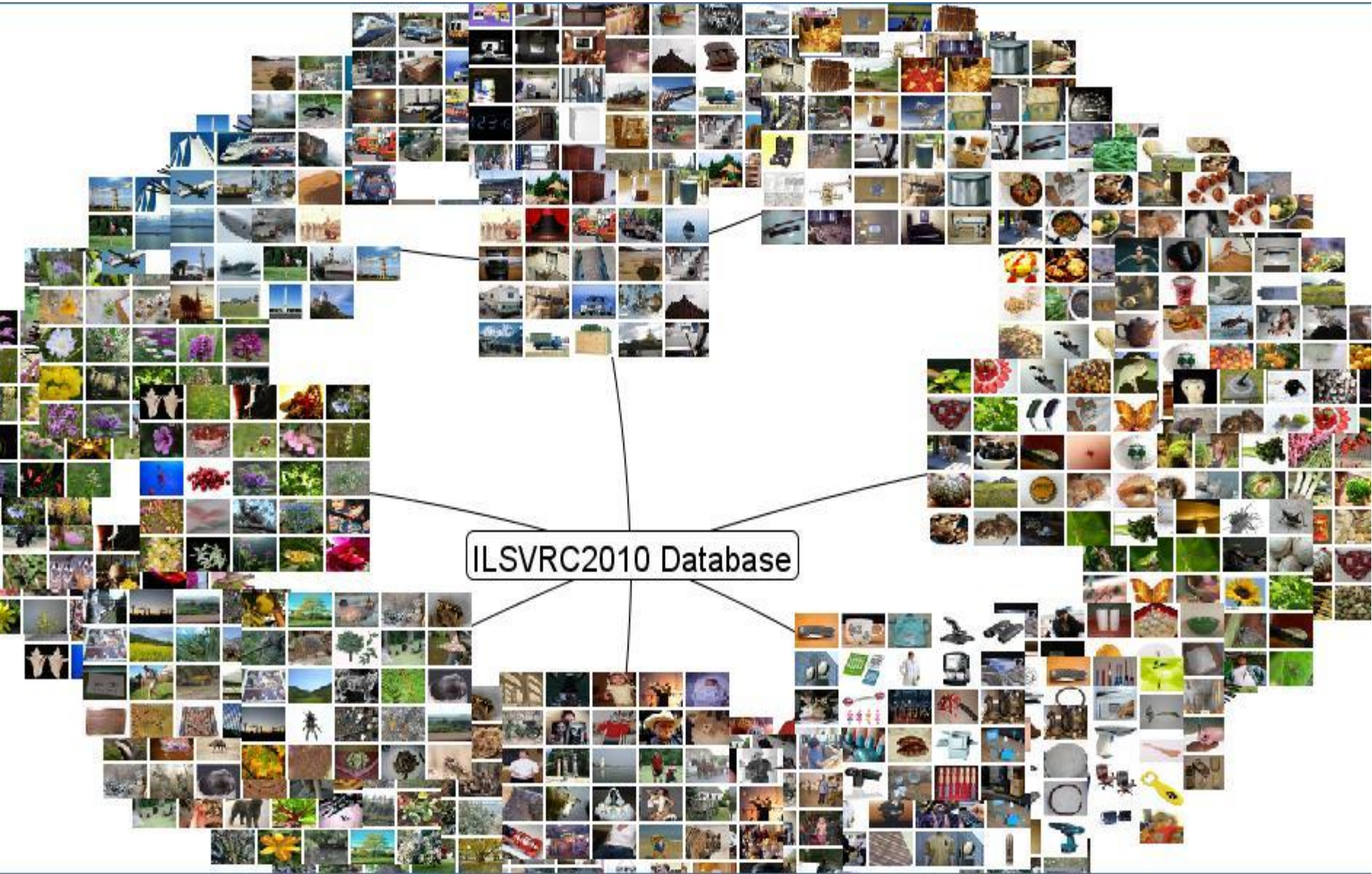
# 4. Visual Tree Construction: Hierarchical Clustering



# 4. Visual Tree Construction: Hierarchical Clustering



# 4. Visual Tree Construction: Hierarchical Clustering



# 4. Visual Tree Construction: Hierarchical Clustering

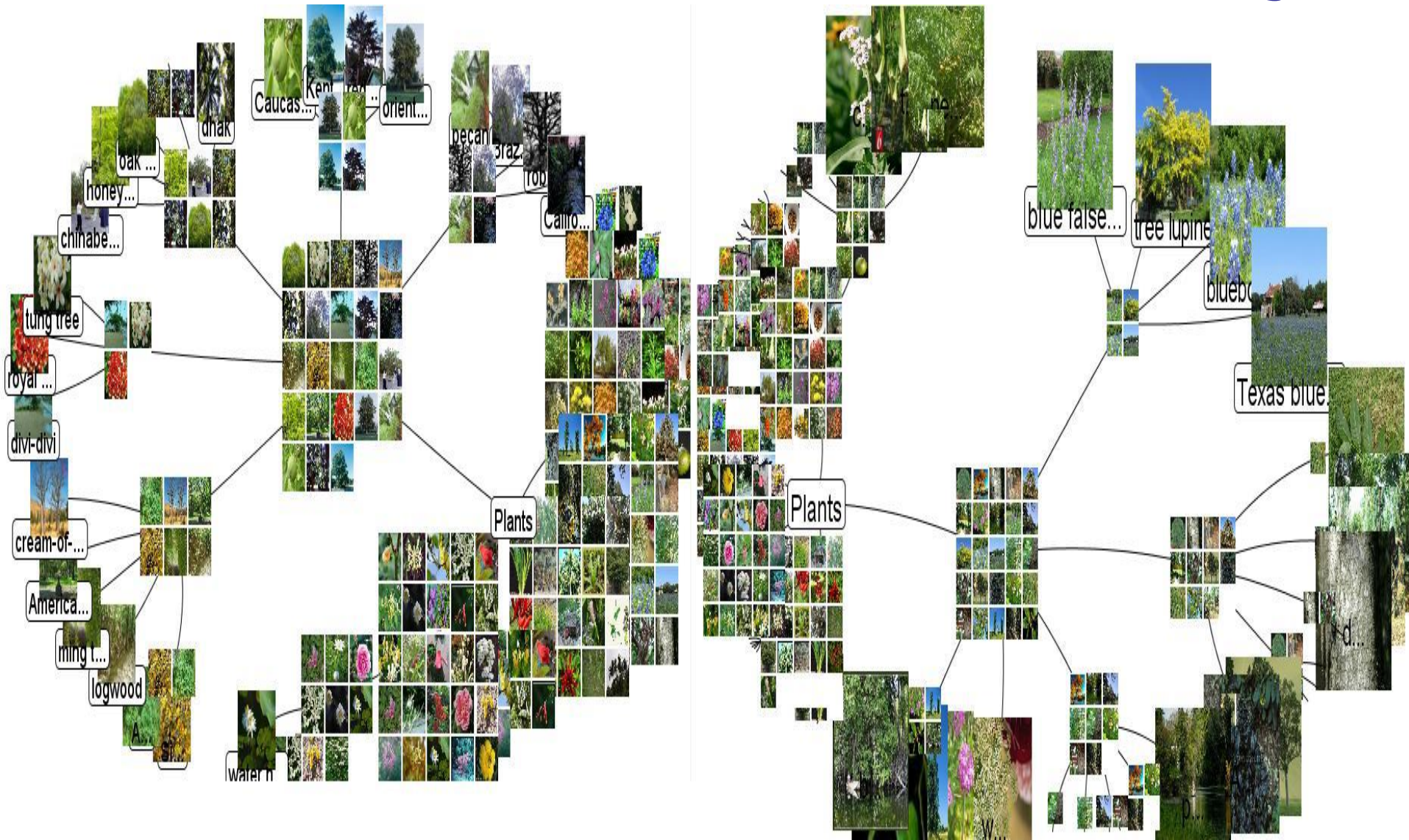


(a)

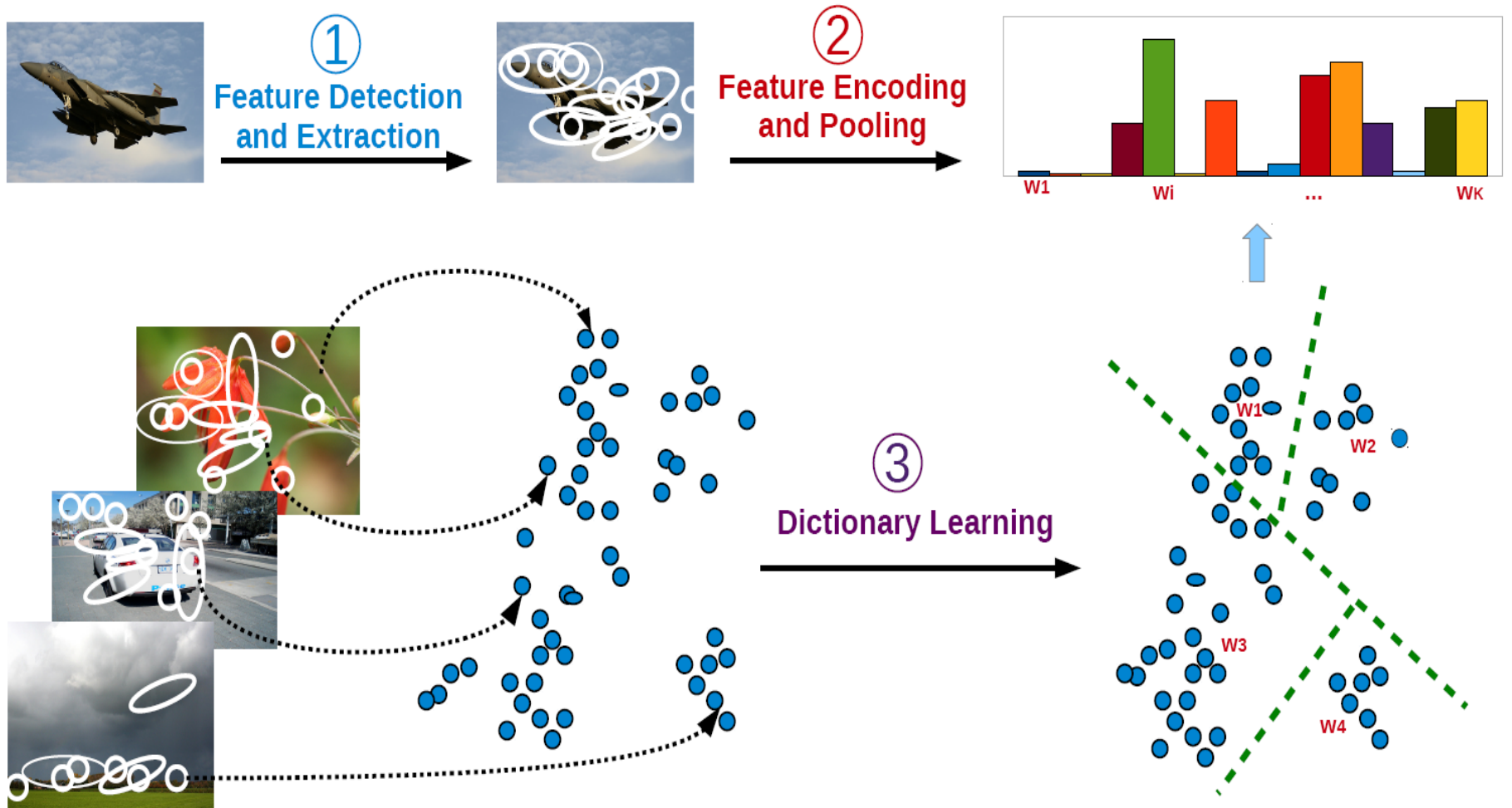


(b)

# 4. Visual Tree Construction: Hierarchical Clustering



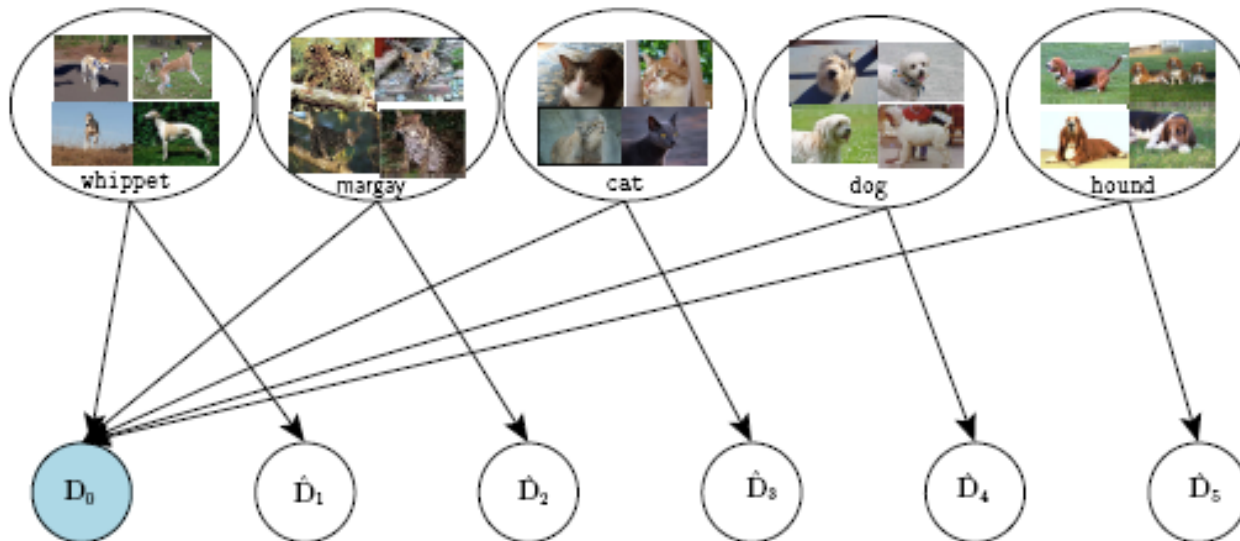
# Bag-of-Words (BoW)





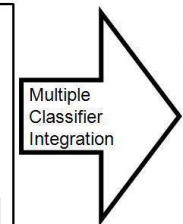
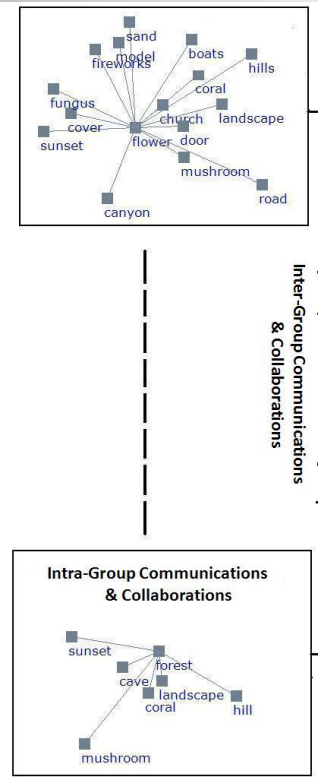
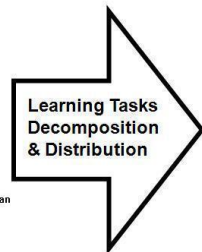
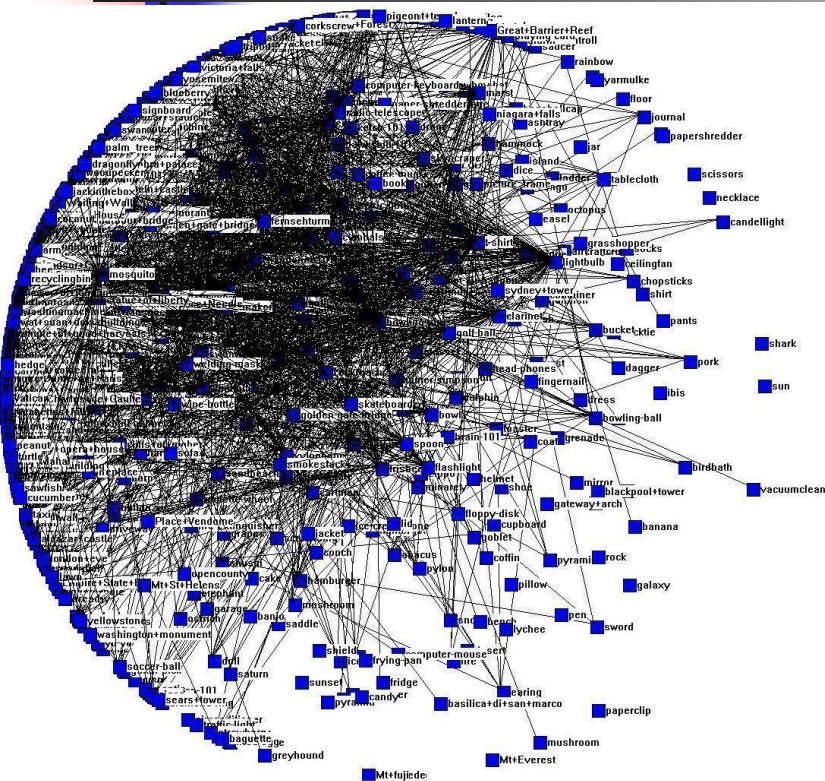
## 5. Joint Dictionary Learning for Discriminative Image Representation

- To distinguish visually-similar categories, dictionaries with strong discrimination is critical
- Joint dictionary learning



CVPR 2012  
TPAMI2013

# 7. Large-Scale Classifier Training



Large Number of Inter-Related Classifiers For Object and Concept Detection

(a) Concept Network for Task and Communication Organization

(b) Training Groups and Inter-Group Communication

(c) Classifier Integration



## 7. Large-Scale Classifier Training

---

### ■ Inference Model Selection for Classifier Training

$$f_{C_j}(x) = W_j^{tr} \Phi_j(x) + \sum_{C_t \in \Theta_j} \gamma_t \cdot V_t^{tr} \Phi_t(x), \quad \sum_{C_t \in \Theta_j} \gamma_t = 1$$

If the given image concept  $C_j$  is visually-related with the image concept  $C_t$  (i.e.,  $C_j$  is linked with  $C_t$  on the visual concept network),  $V_t \neq 0$ . If the given image concept  $C_j$  is visually-irrelevant with the image concept  $C_t$  (i.e.,  $C_j$  is not linked with  $C_t$  on the visual concept network),  $V_t = 0$ .

$$J = \frac{1}{2} (\|W_j\|^2 + \sum_{t=1}^{|\Theta_j|} \lambda_t \|V_t\|^2) + \rho_0 \sum_{t=1}^{|\Theta_j|} \sum_{i=1}^{n_j} \xi_{ti} + \sum_{t=1}^{|\Theta_j|} \rho_t \sum_{i=1}^{n_t} \eta_{ti}$$



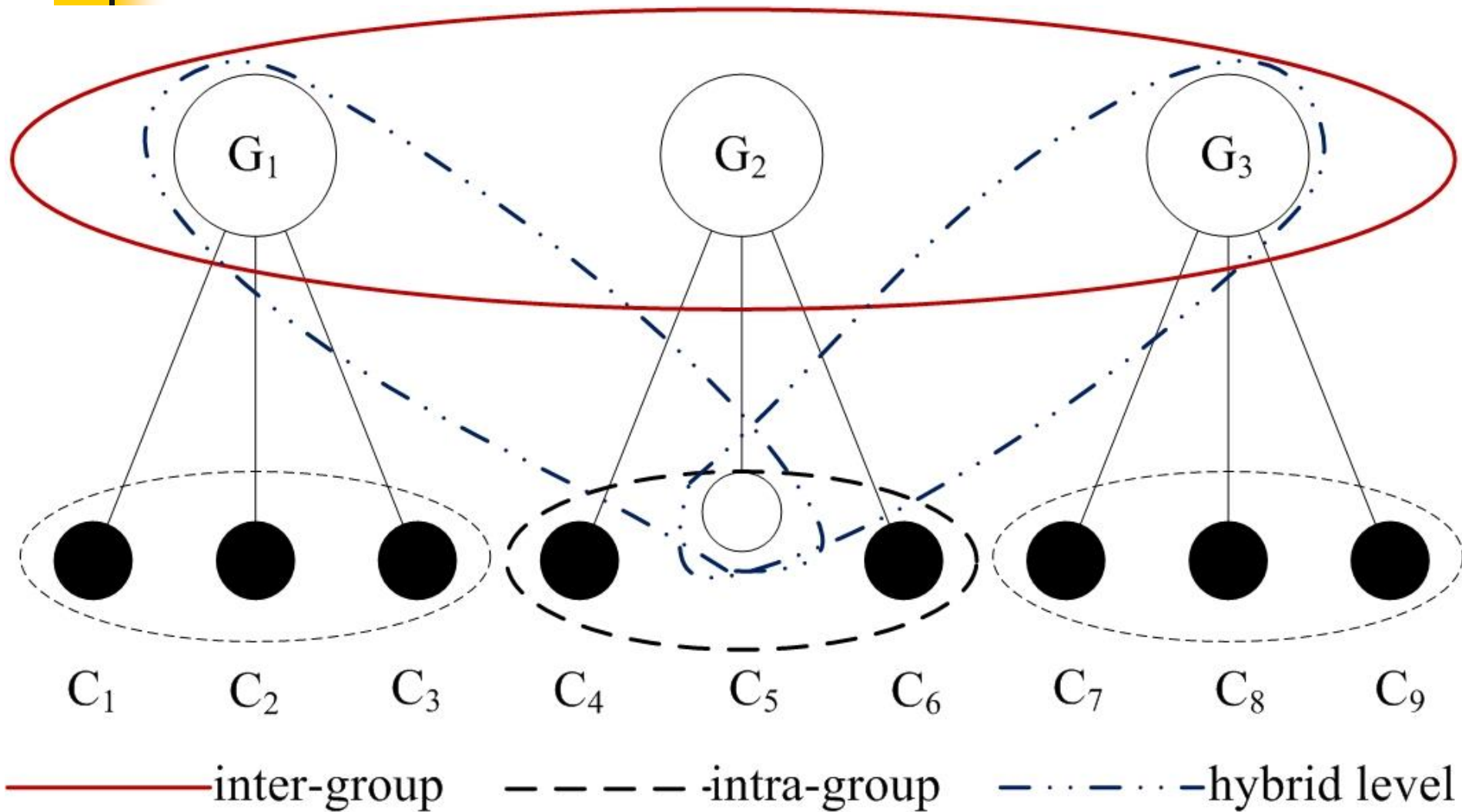
## 7. Large-Scale Classifier Training

---

### ■ Inference Model Selection for Classifier Training

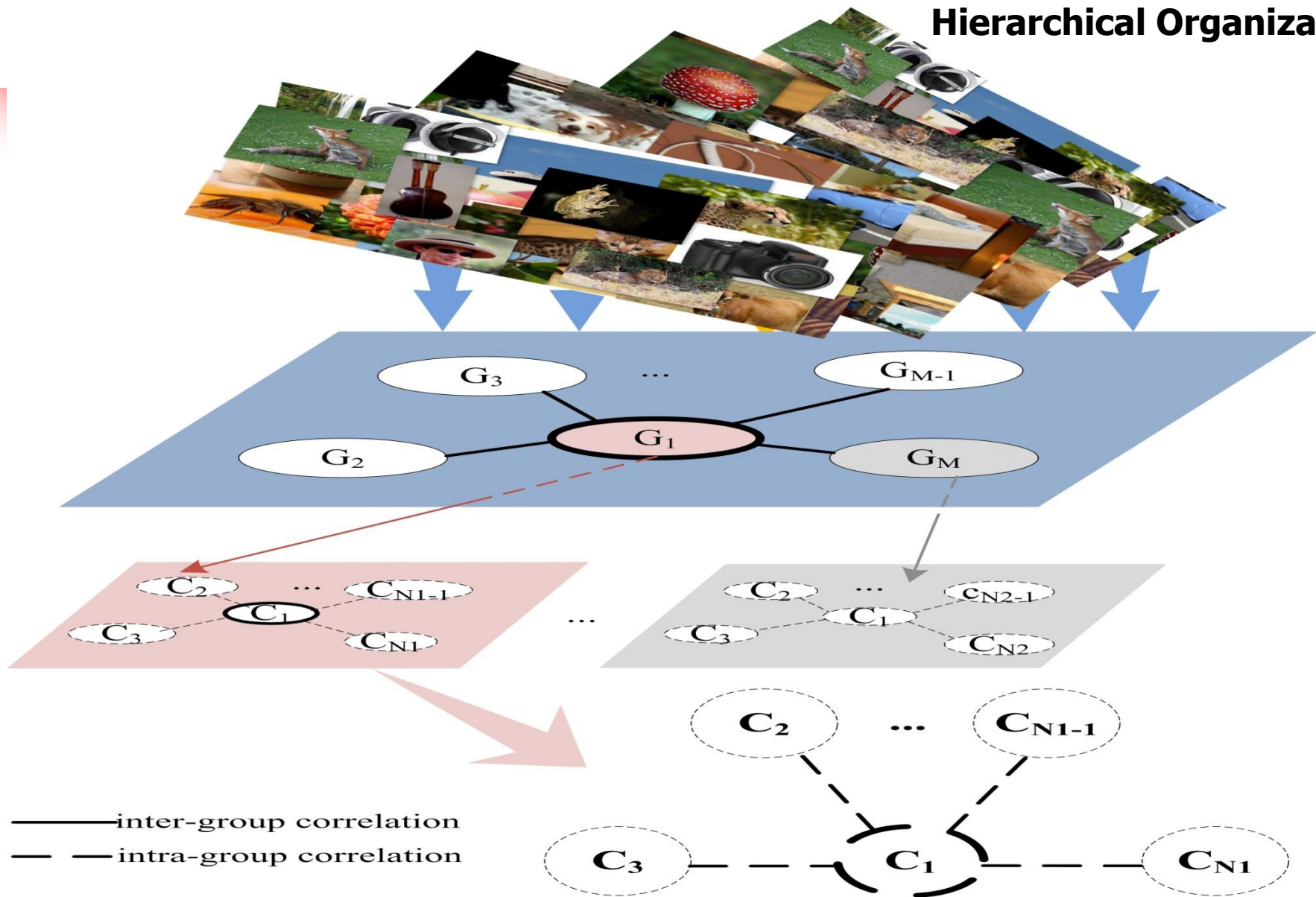
$$f_{C_j}(x) = \sum_{h,t=1}^{|\Theta_j|} \gamma_t \kappa_s(t, h) \left( \sum_{i=1}^{n_j} \beta_{hi} \kappa(x_{ji}, x) - \sum_{i=1}^{n_h} \overline{\beta_{hi}} \kappa(x_{hi}, x) \right) \\ + \sum_{t=1}^{|\Theta_j|} \frac{\gamma_t}{\lambda_t} \left( \sum_{i=1}^{n_j} \beta_{ti} \kappa(x_{ji}, x) - \sum_{i=1}^{n_t} \overline{\beta_{ti}} \kappa(x_{ti}, x) \right)$$

## 7. Large-Scale Classifier Training



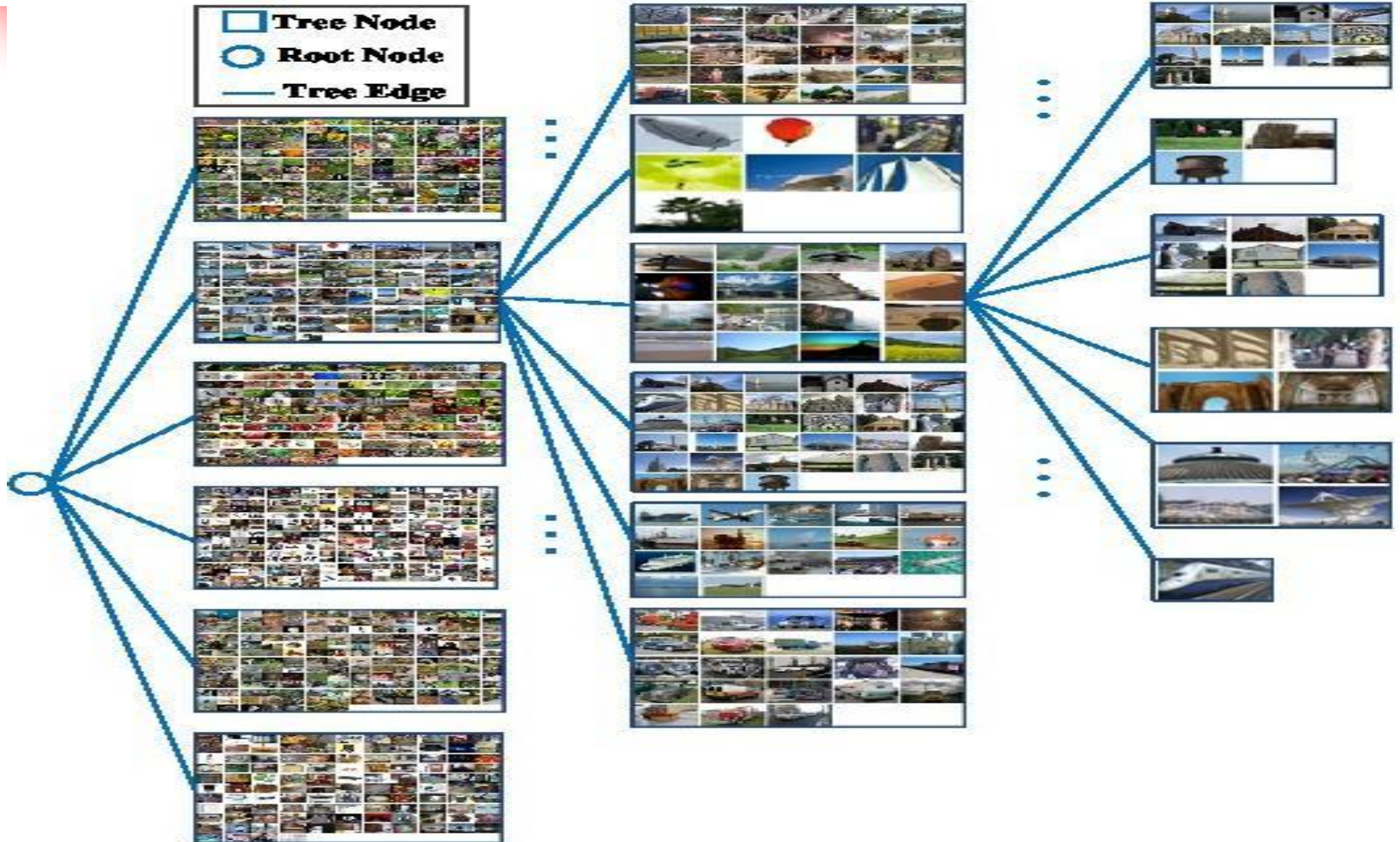
# 7. Large-Scale Classifier Training

## Hierarchical Organization



# 7. Large-Scale Classifier Training

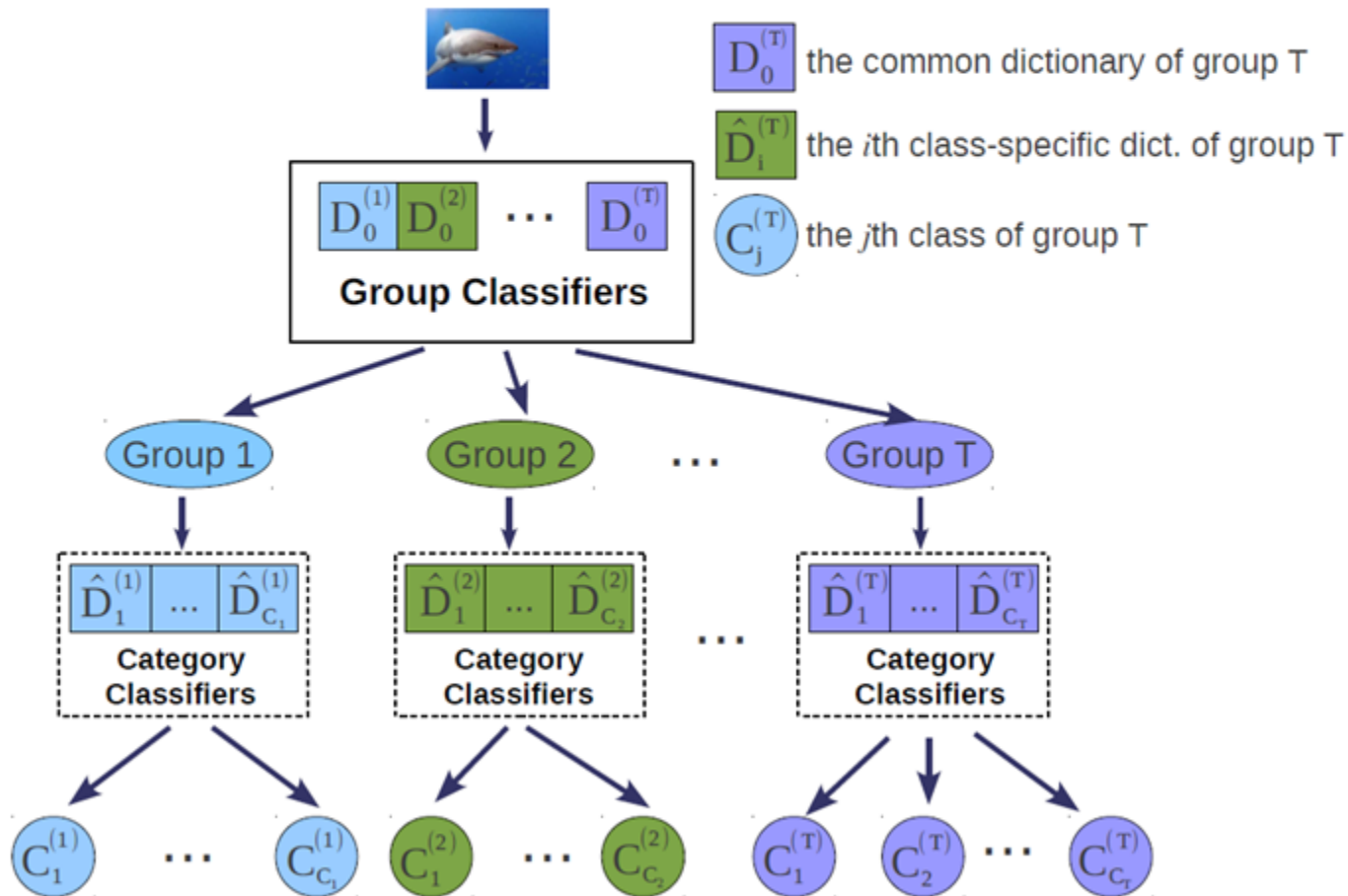
## Hierarchical Organization



# 7. Large-Scale Classifier Training

## Hierarchical Organization

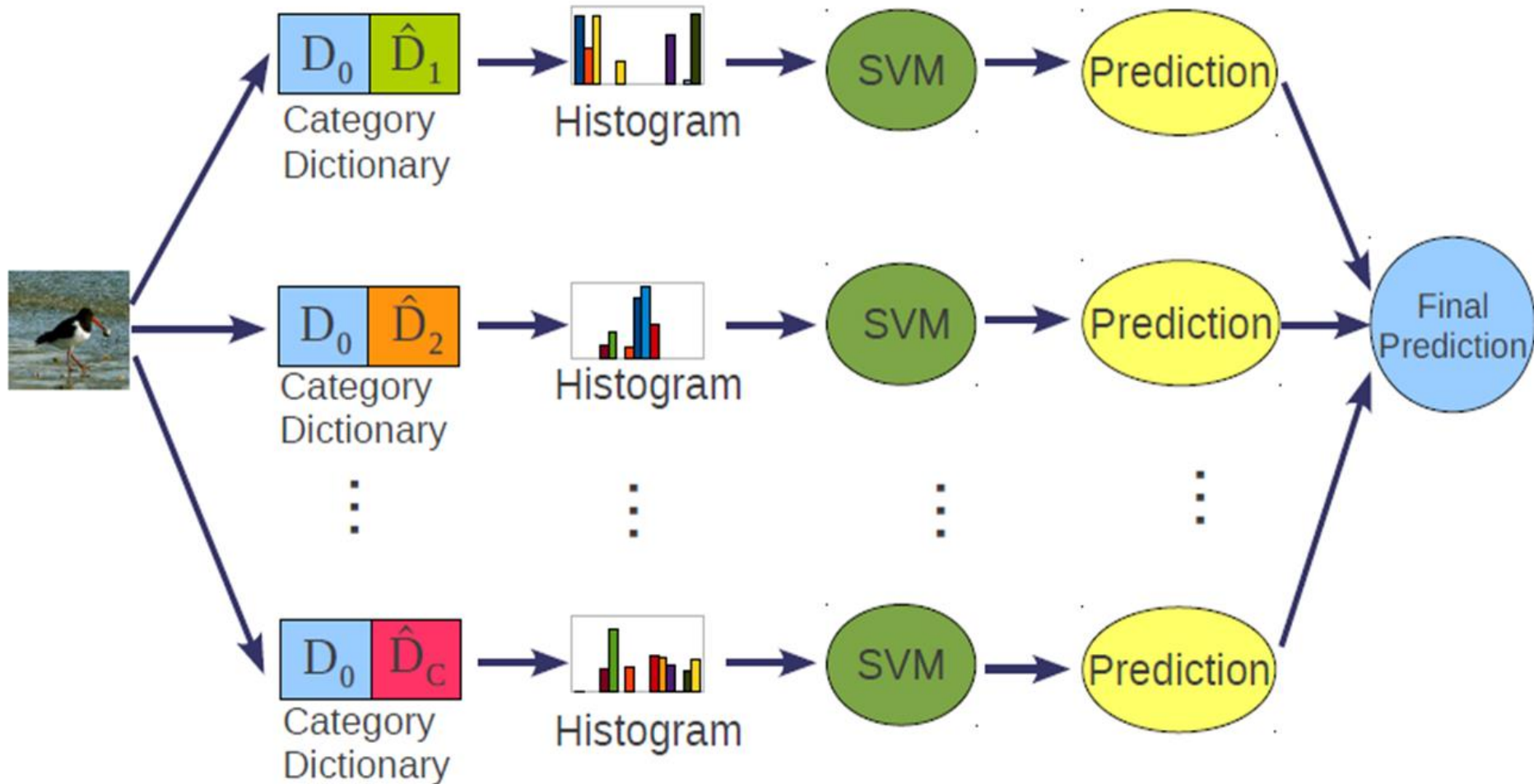
### Hierarchical Classification Scheme





# 7. Large-Scale Classifier Training

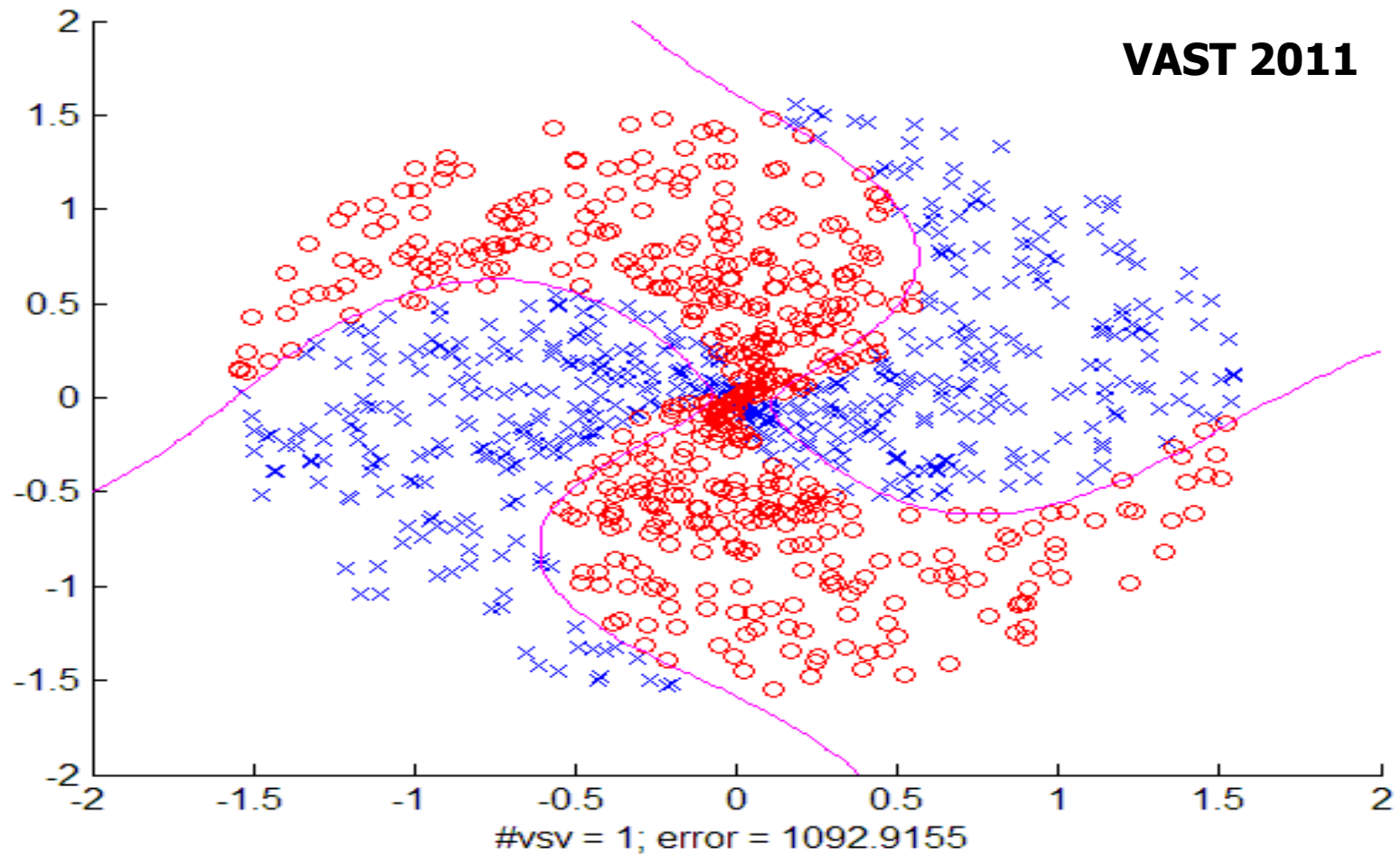
**Flat Organization**





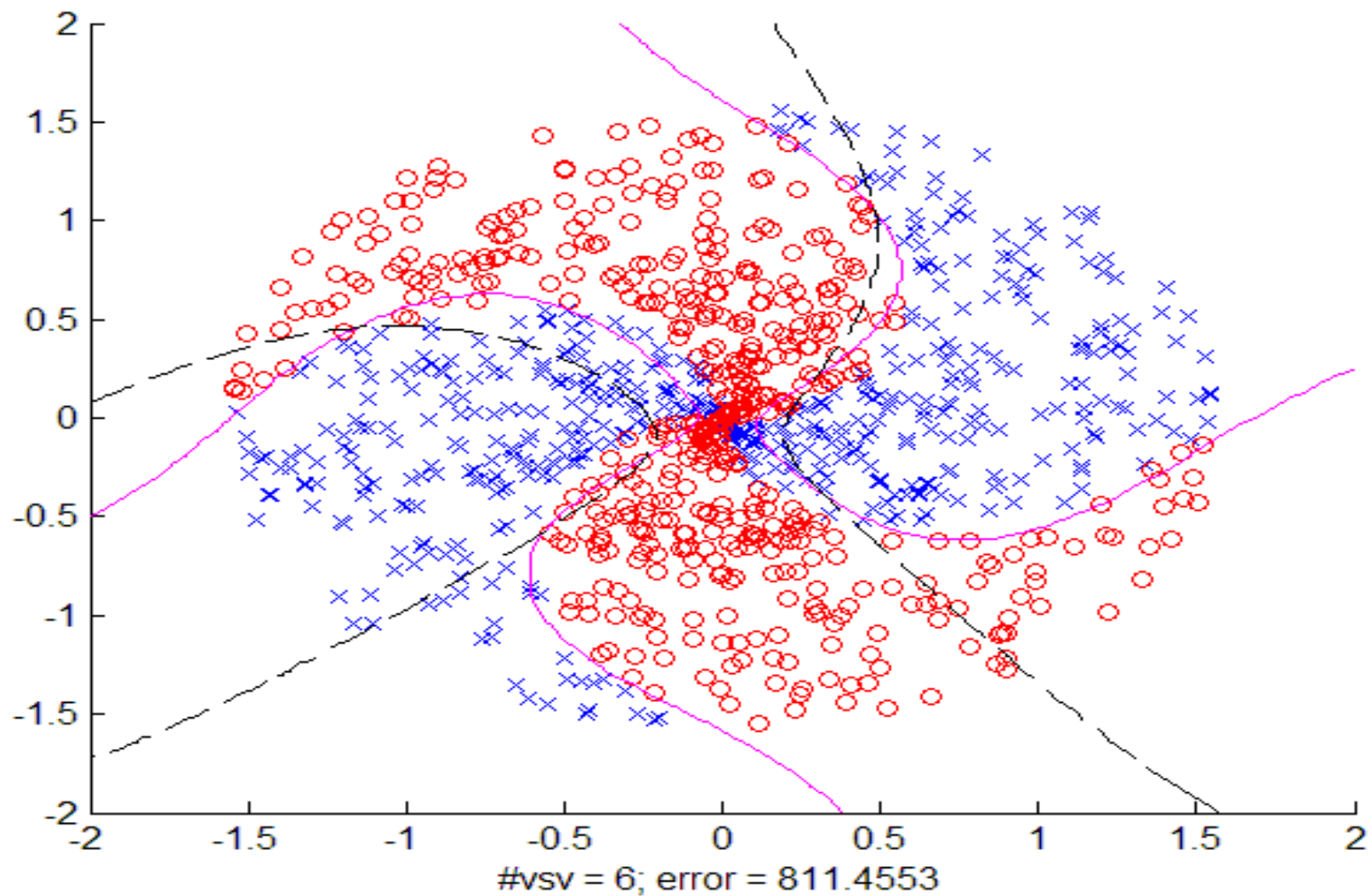
## 8. Interactive Classifier Assessment

**VAST 2011**

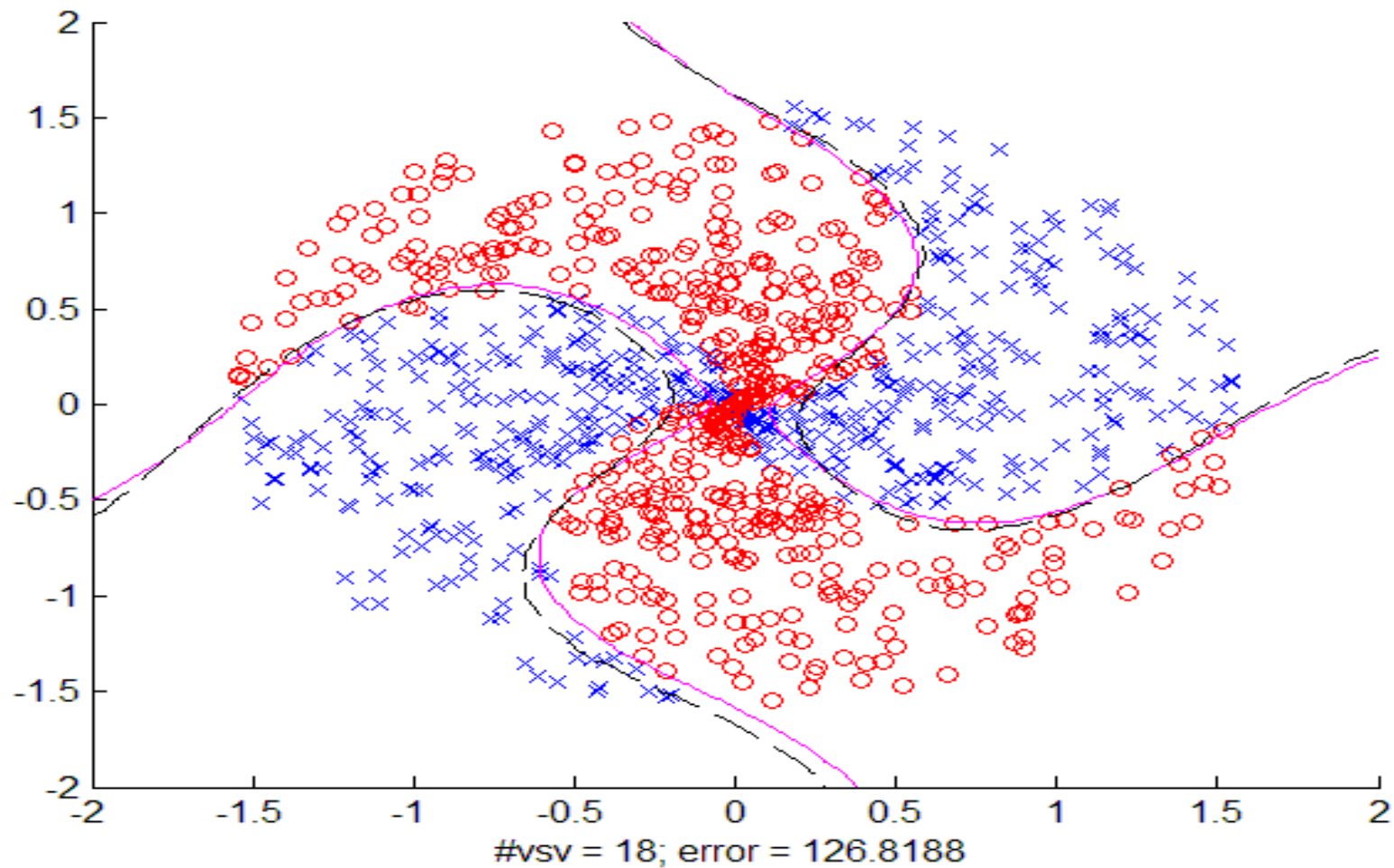




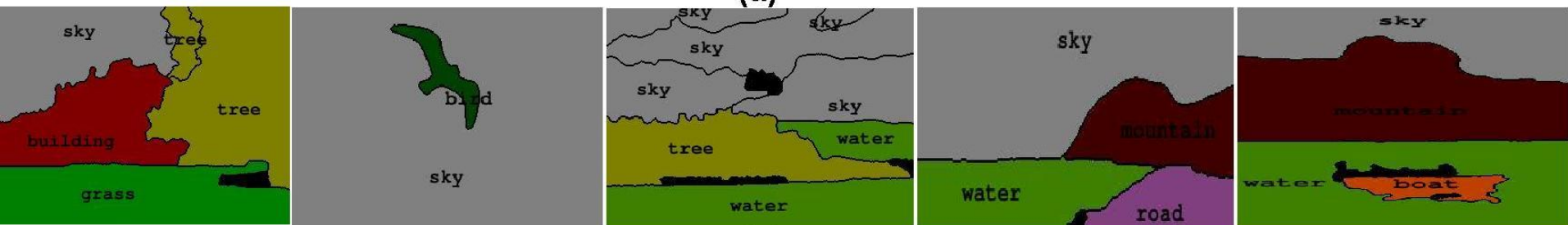
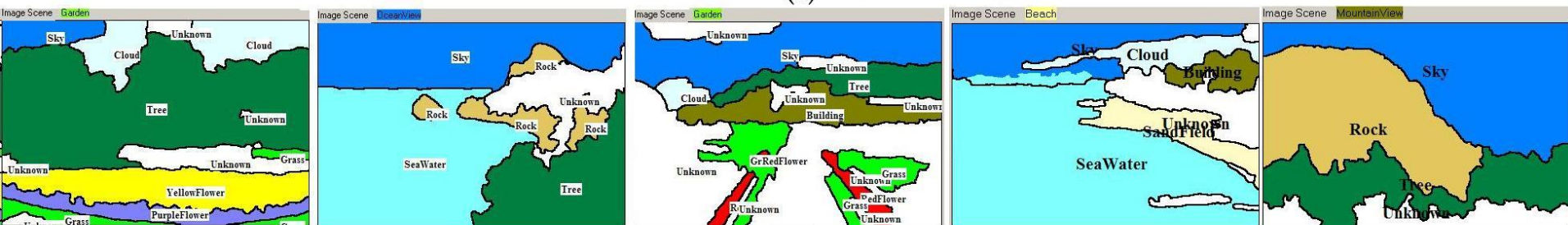
## 8. Interactive Classifier Assessment



## 8. Interactive Classifier Assessment



# 9. Some Experimental Results



# 9. Some Experimental Results



(a)

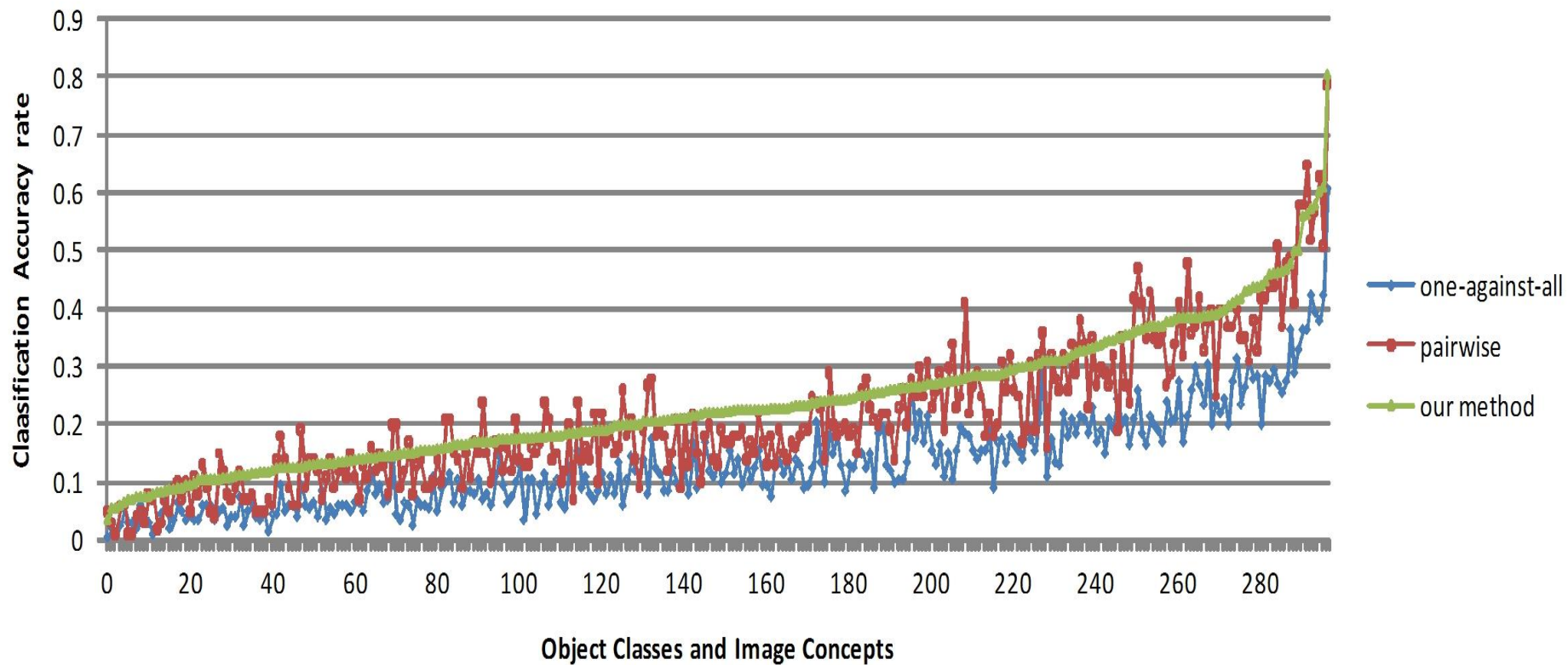


(b)

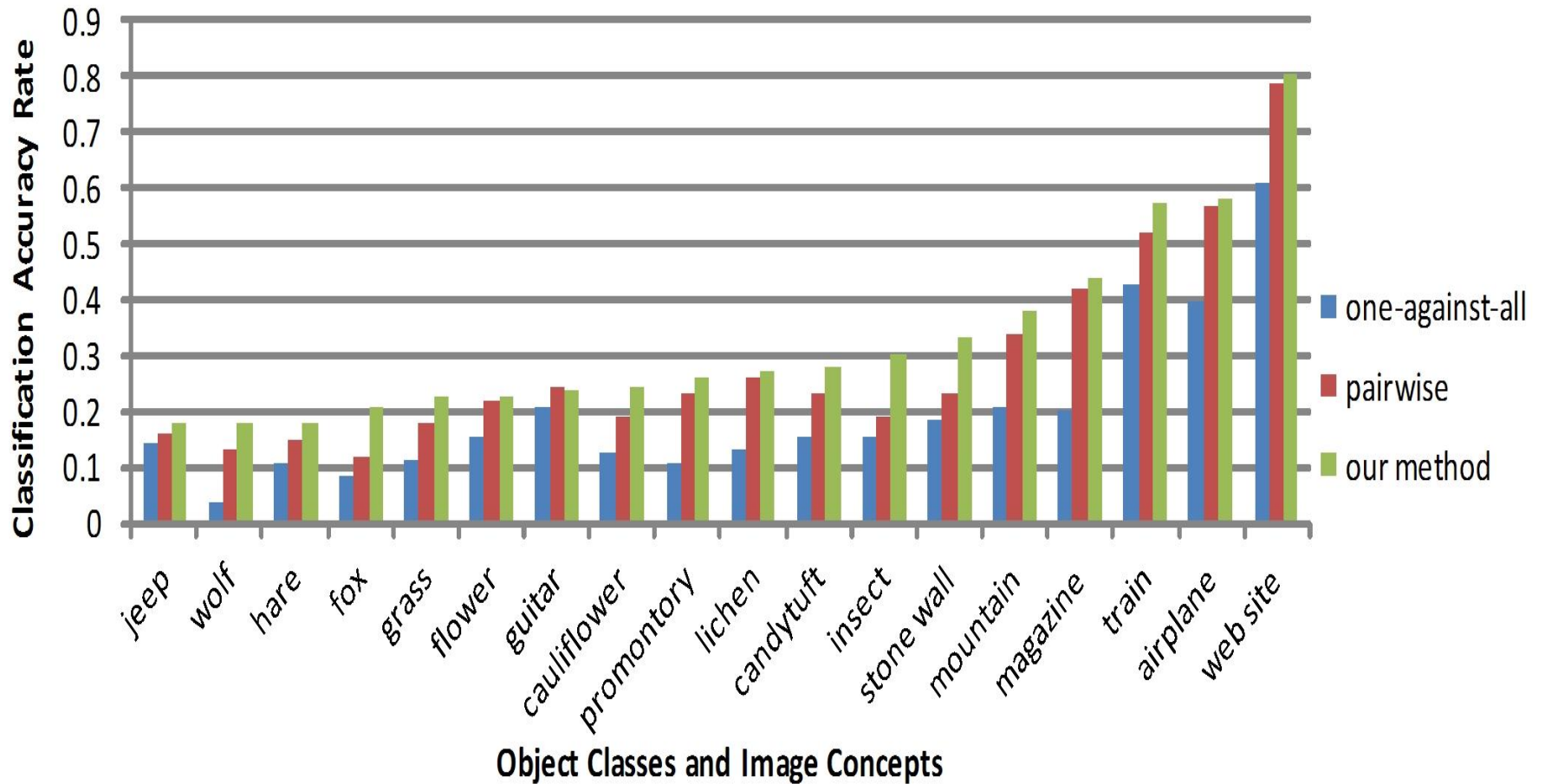


(c)

## 9. Some Experimental Results

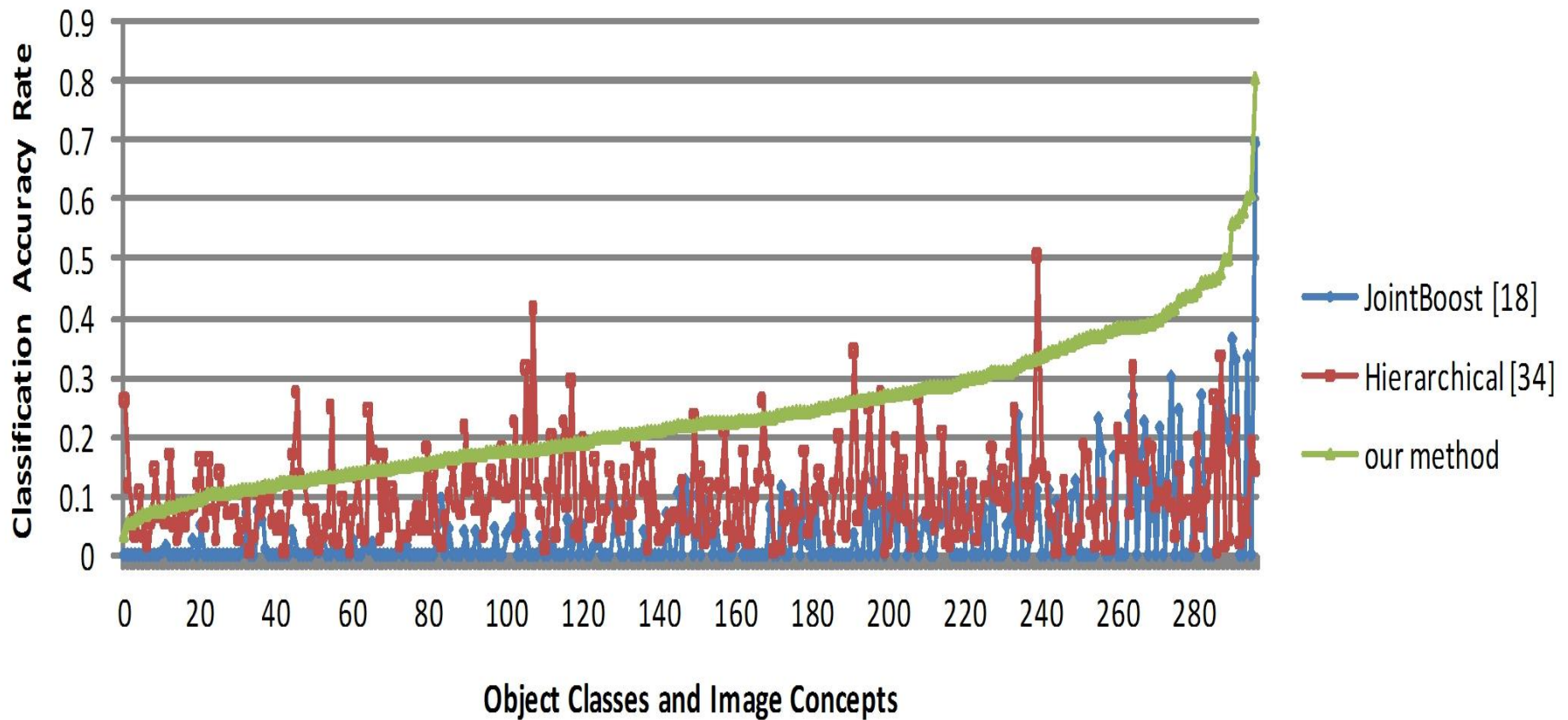


## 9. Some Experimental Results

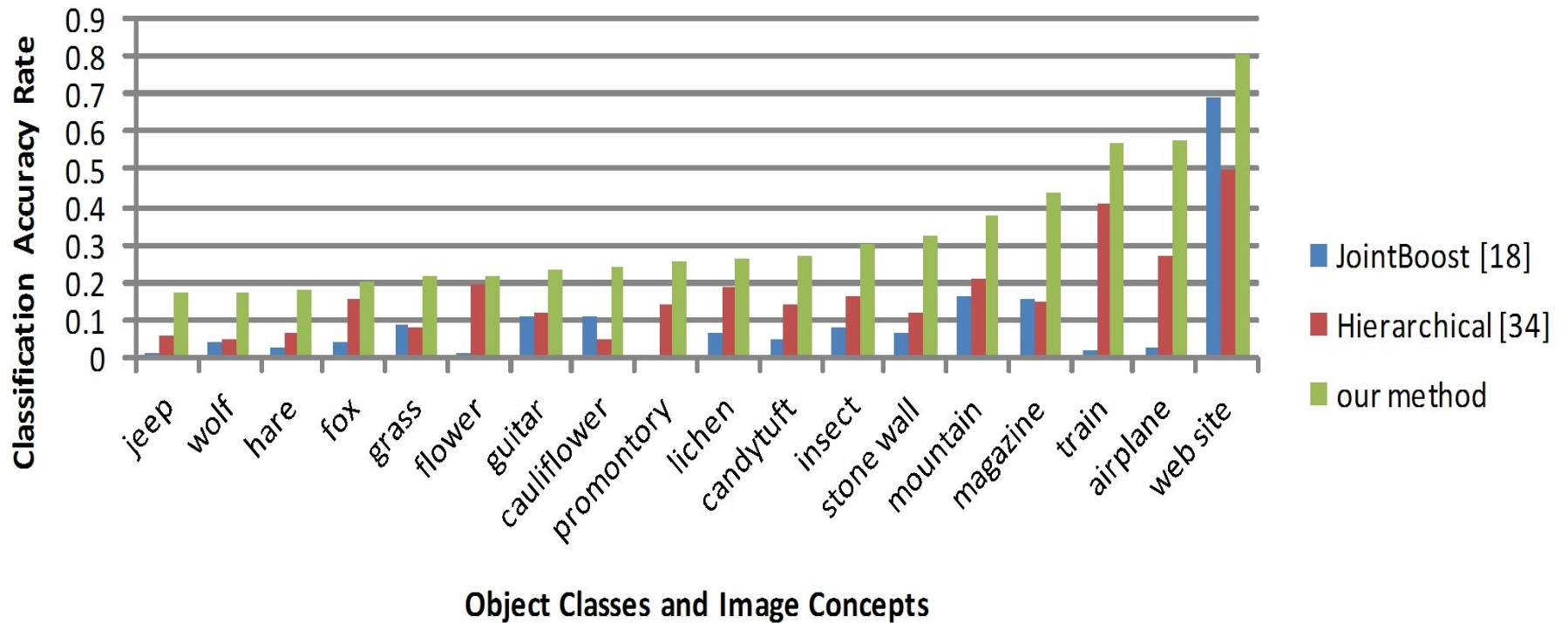




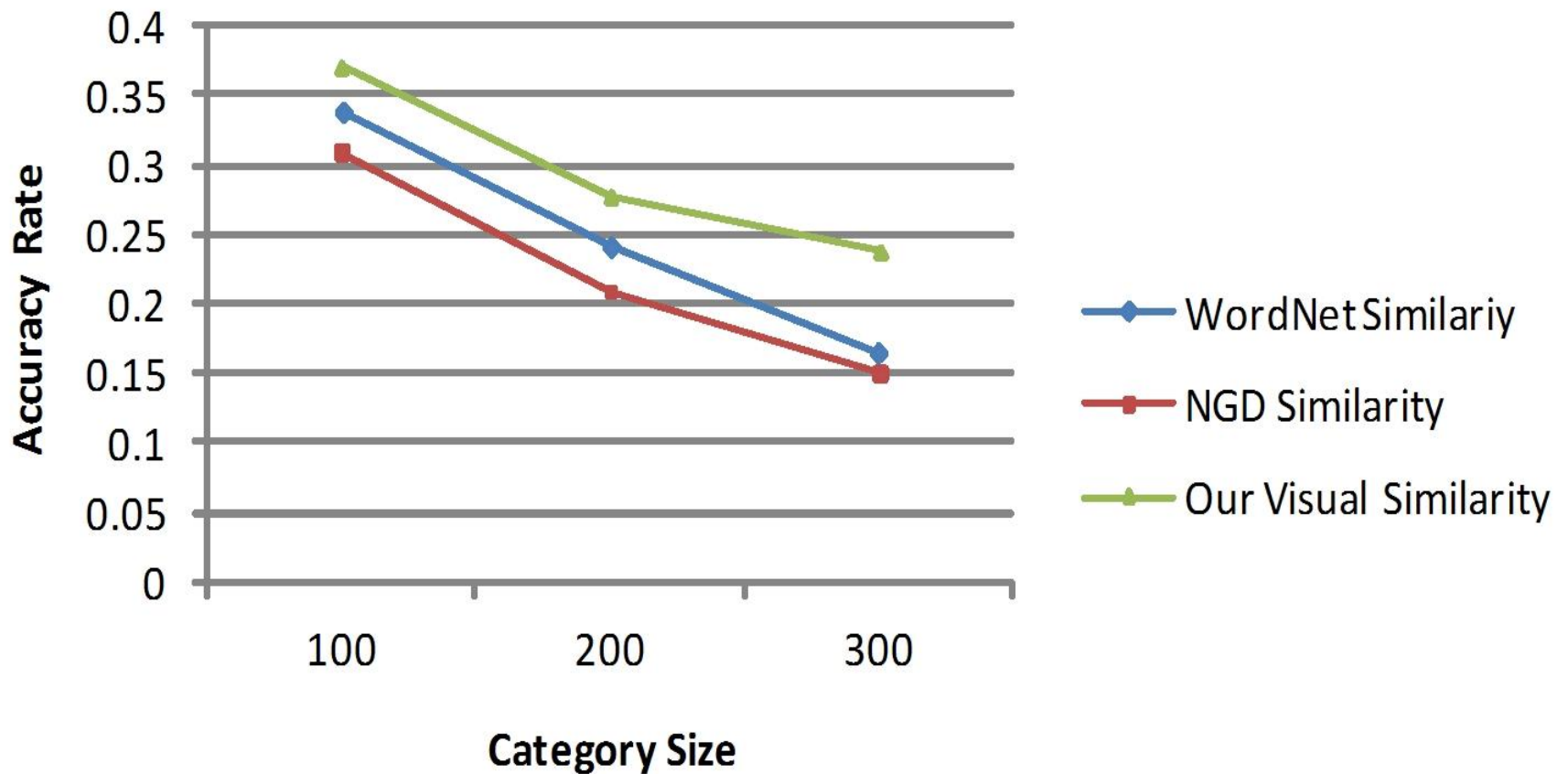
## 9. Some Experimental Results



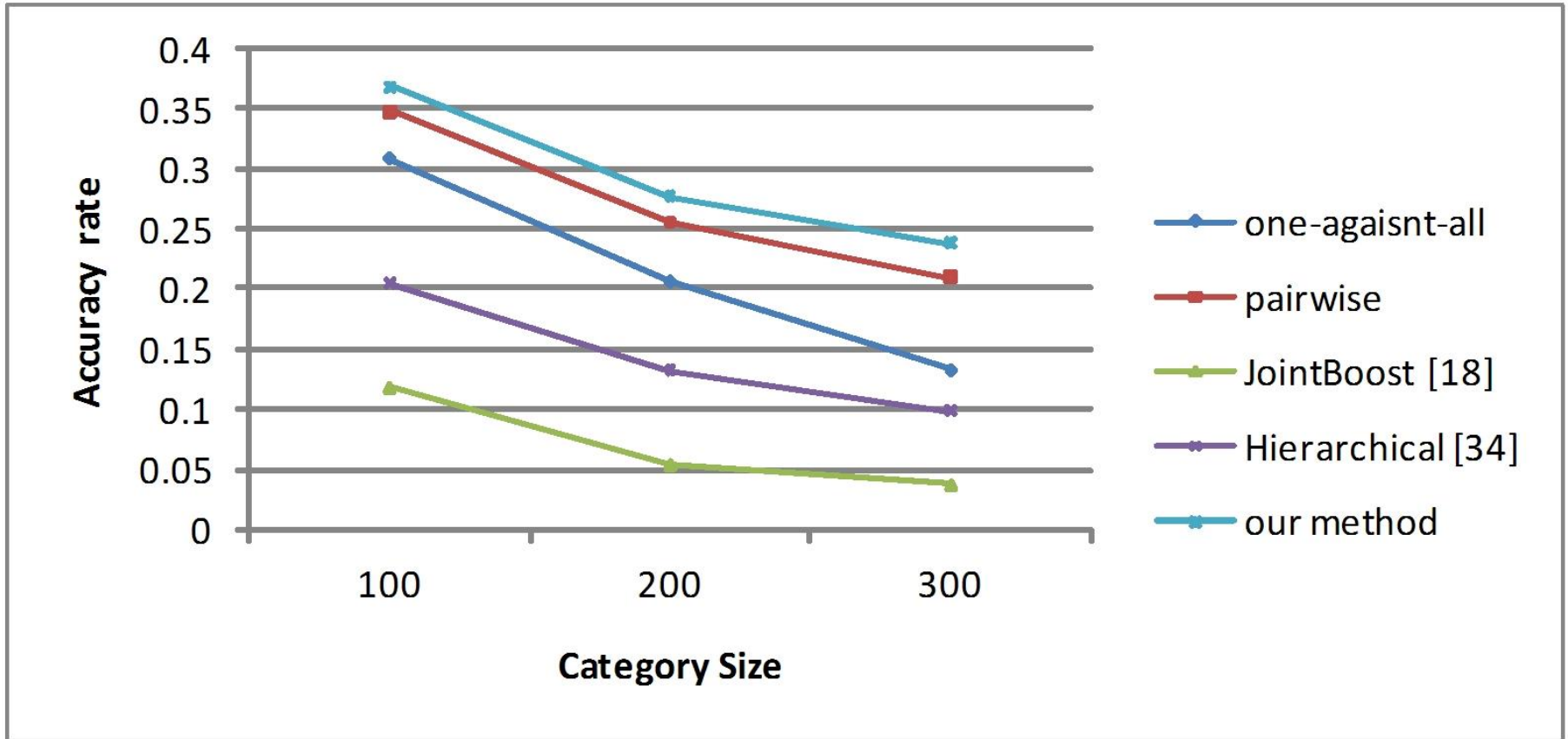
## 9. Some Experimental Results



## 9. Some Experimental Results

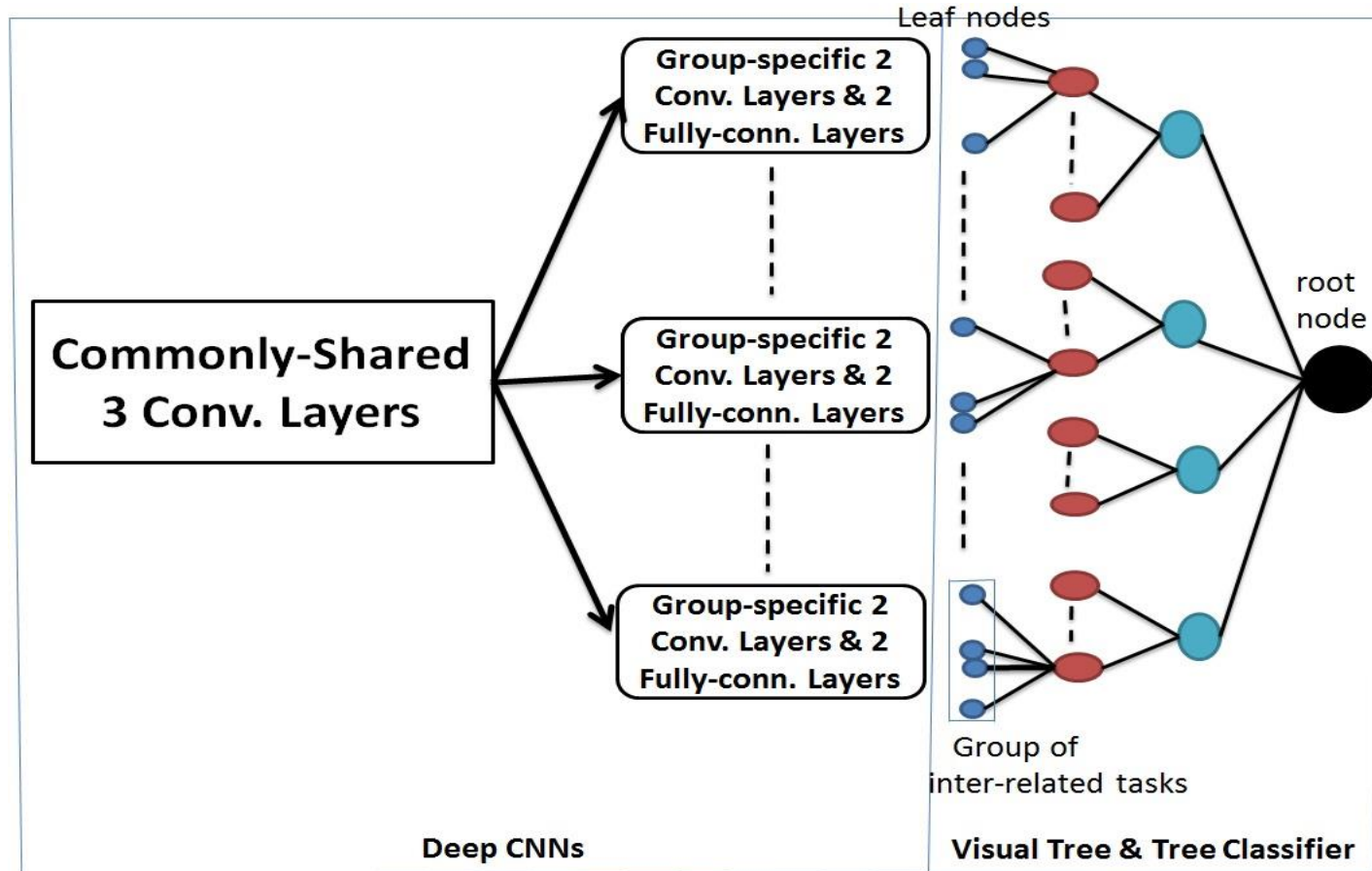


## 9. Some Experimental Results



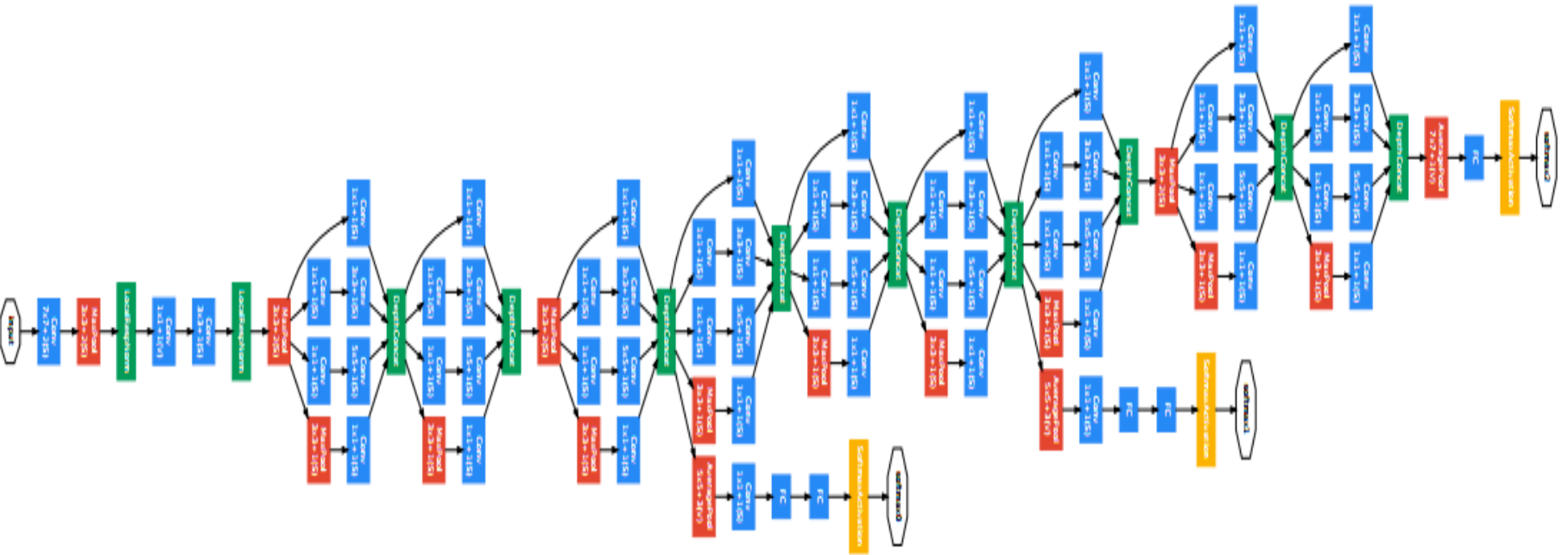
# Hierarchical Deep Multi-Task Learning (HD-MTL) over Visual Tree

## ■ HD-MTL



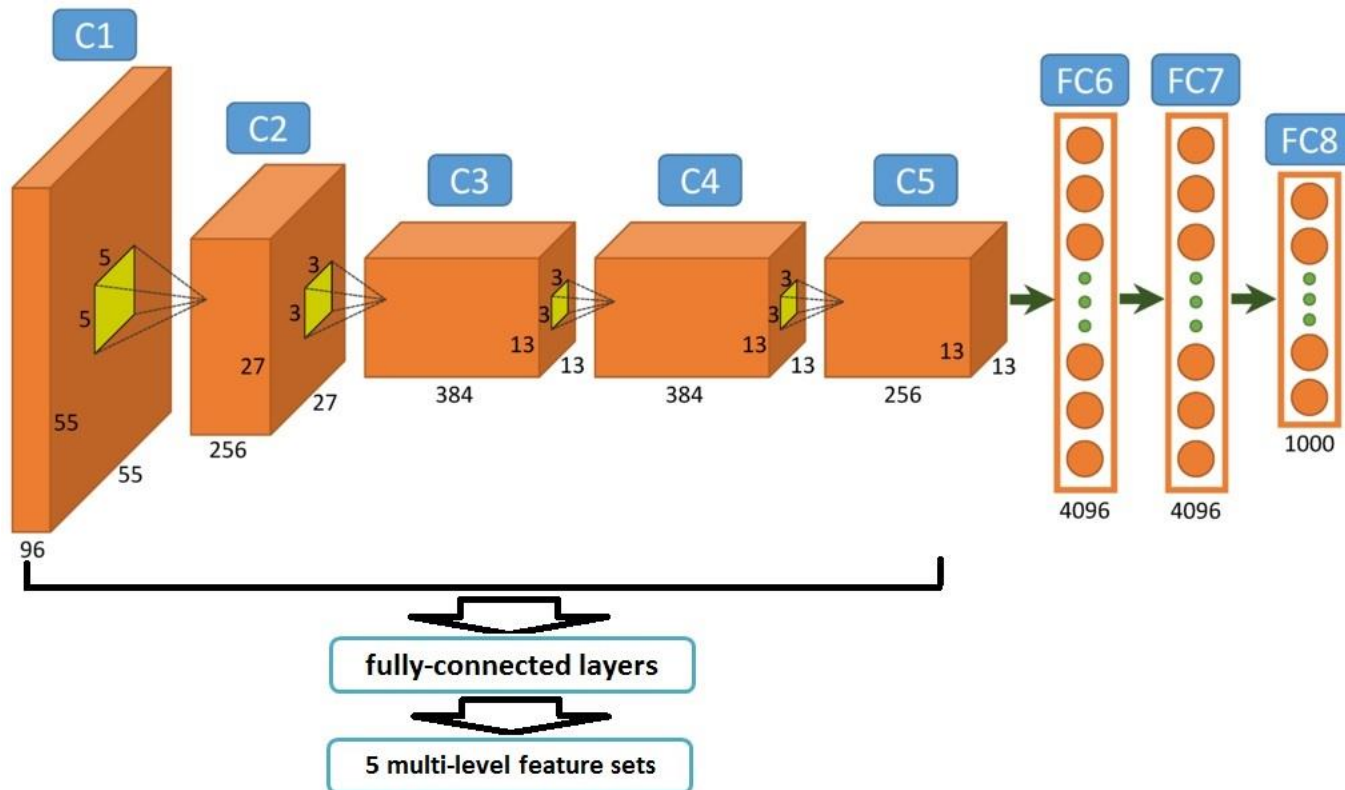
# 2. Multi-Level Deep Feature Extraction

- GoogleNet**



## 2. Multi-Level Deep Feature Extraction

### ■ Deep CNNs for Feature Extraction

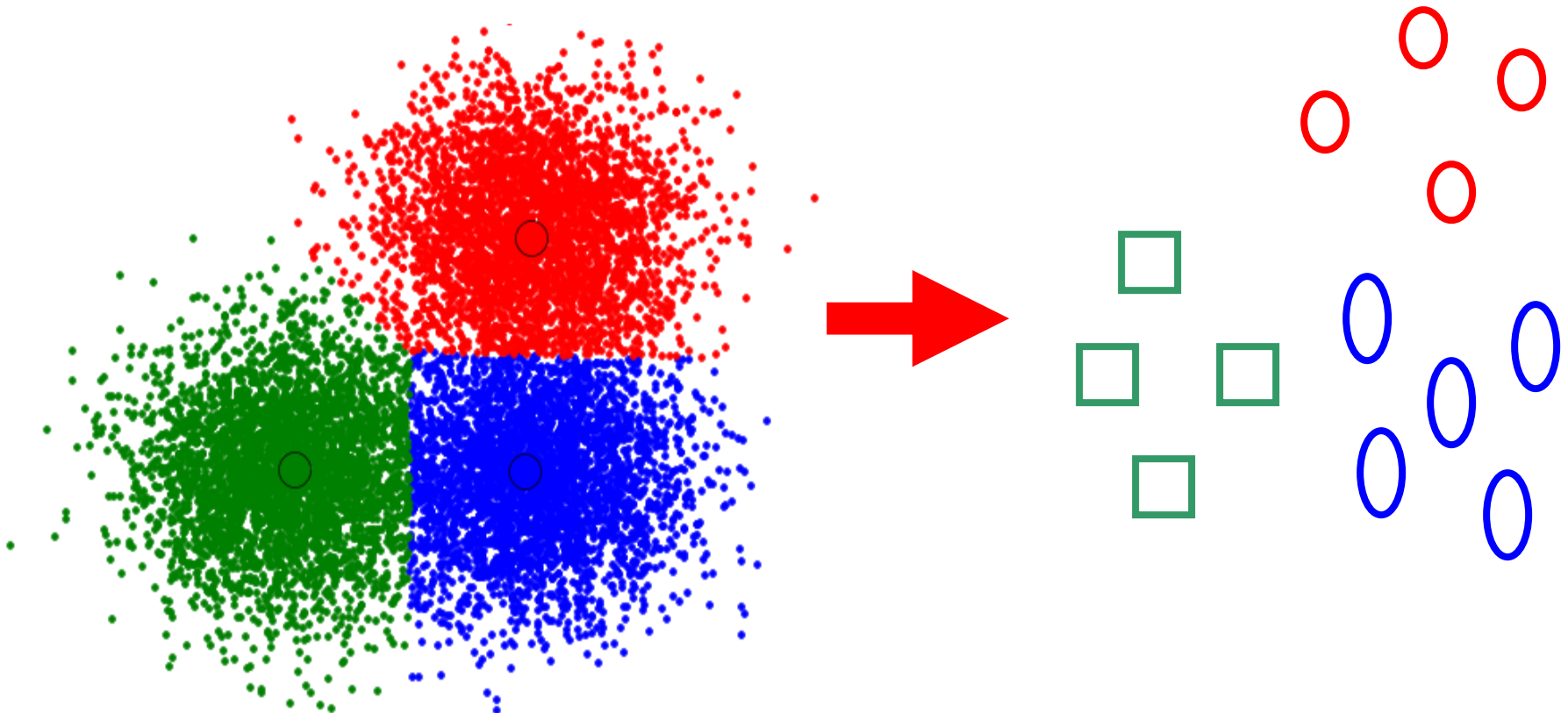




## 3. Visual Tree Construction

---

- **Object Class Representation**







## 4. Visual Tree Construction: Hierarchical Clustering

---

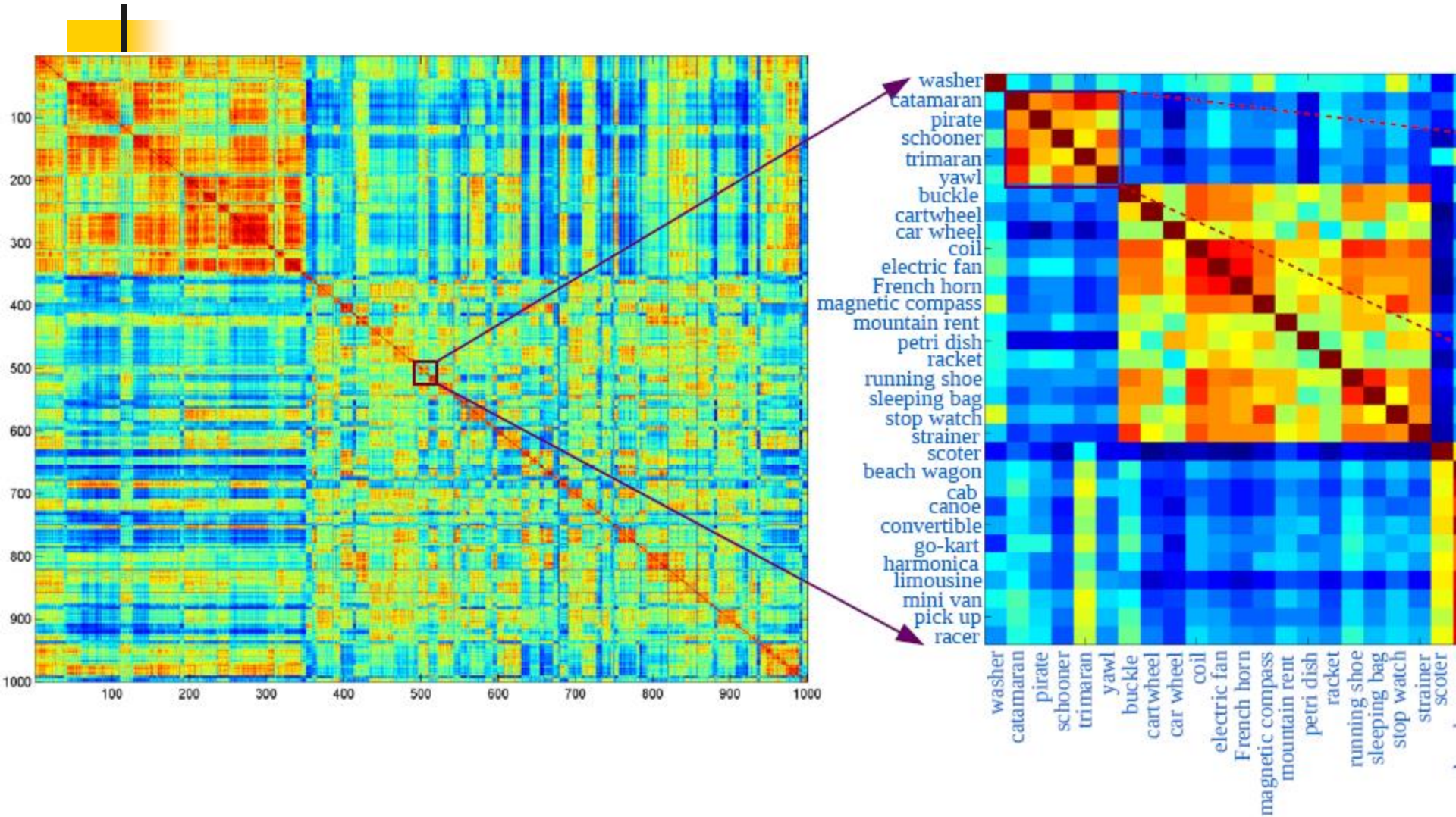
### ■ Feature Subset Selection

$$F_{best}^c = \max \left\{ \Phi_t^c = \frac{1}{\sum_{i=1}^M \sum_{j=1}^M \kappa^t(i, j)}, F_t \in \mathbb{F} \right\}$$

### ■ Node Partitioning

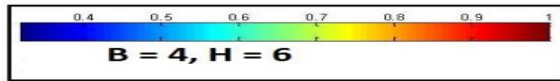
$$\min \left\{ \psi(c, B) = \sum_{l=1}^B \frac{\sum_{i \in G_l} \sum_{j \in G^c / G_l} \kappa_t(i, j)}{\sum_{i \in G_l} \sum_{j \in G_l} \kappa_t(i, j)} \right\}$$

# 4. Visual Tree Construction: Hierarchical Clustering

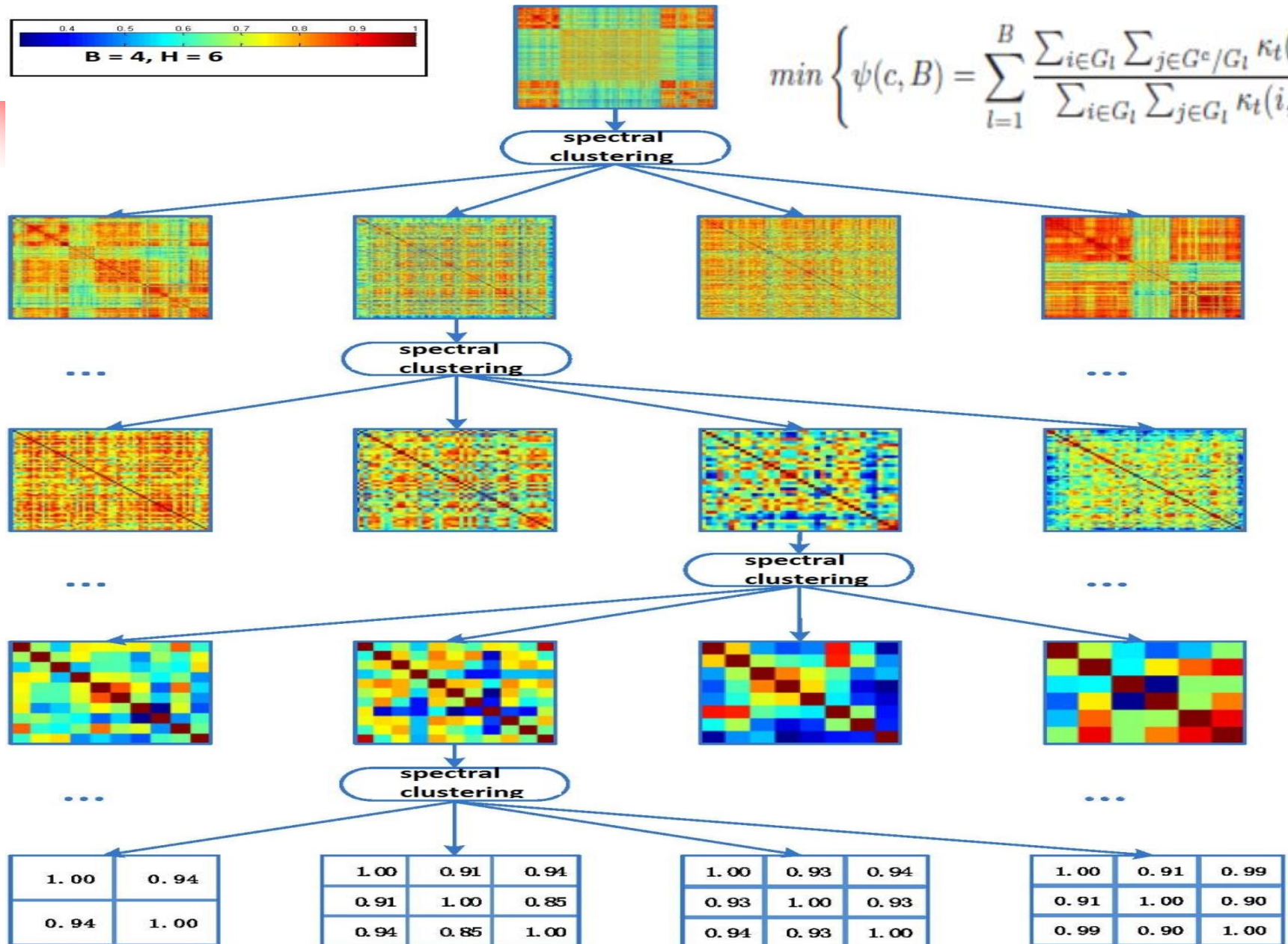


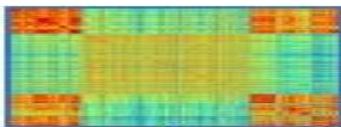
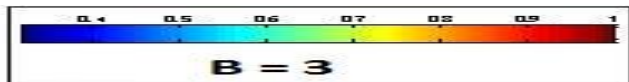
Result is based on ImageNet data set of 1000 categories

# 4. Visual Tree Construction: Hierarchical Clustering

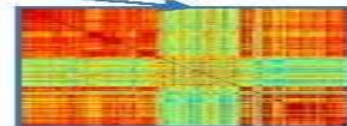
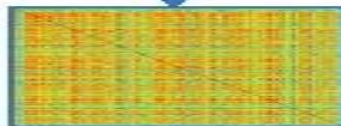
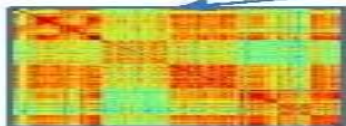


$$\min \left\{ \psi(c, B) = \sum_{l=1}^B \frac{\sum_{i \in G_l} \sum_{j \in G^c/G_l} \kappa_t(i, j)}{\sum_{i \in G_l} \sum_{j \in G_l} \kappa_t(i, j)} \right\}$$





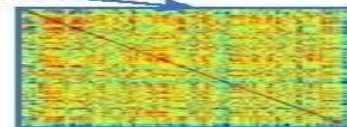
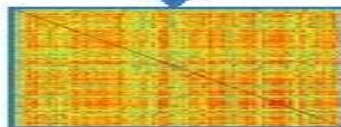
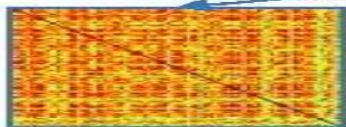
spectral clustering



...

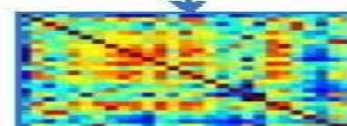
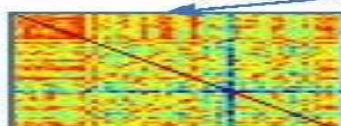
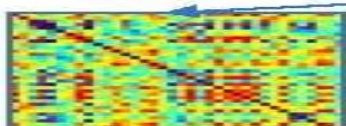
...

spectral clustering



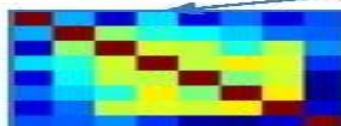
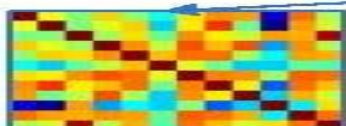
...

spectral clustering



...

spectral clustering



...

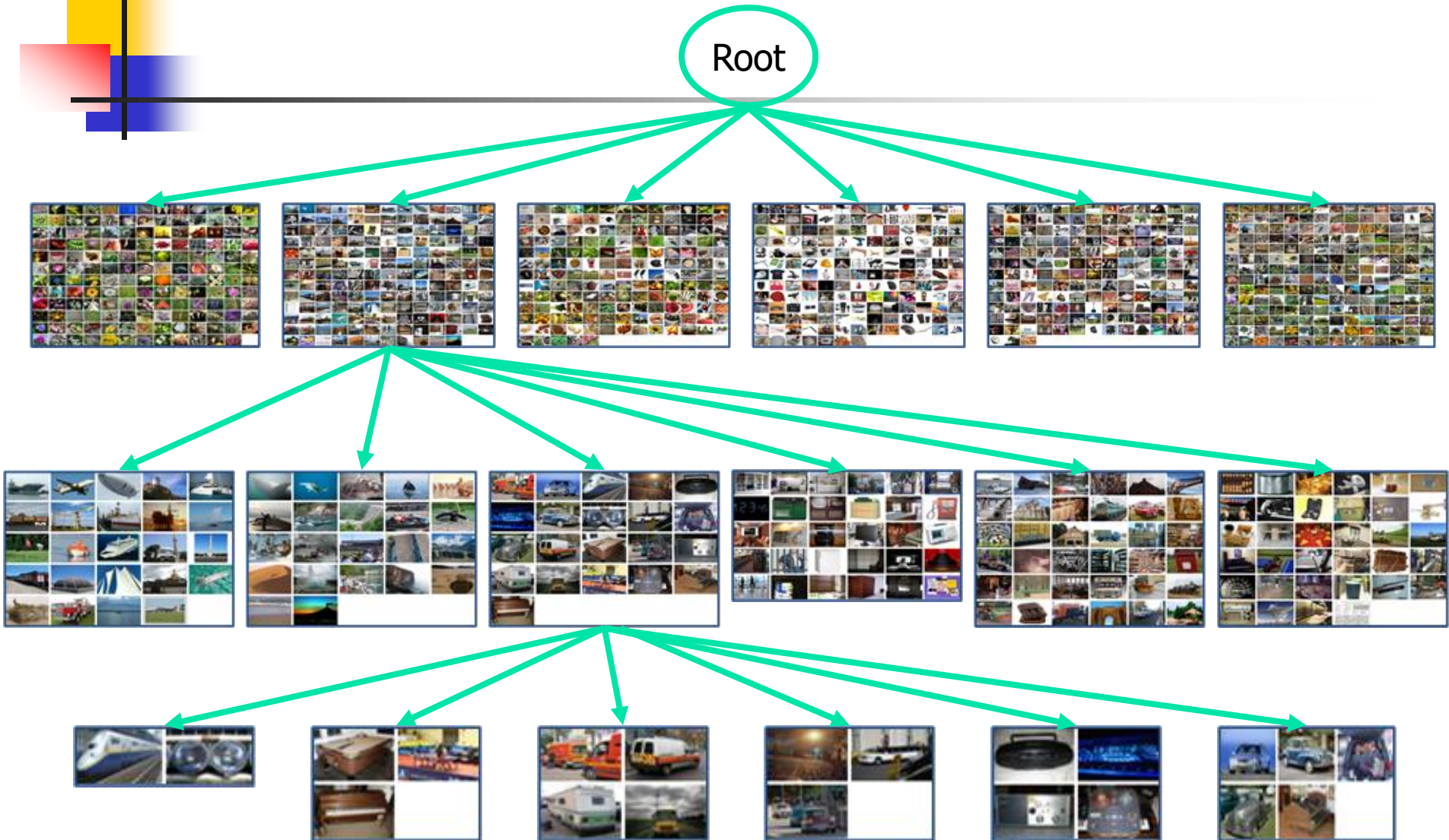
spectral clustering

0.9	0.95	0.9	0.9	0.9
0.95	1.0	0.95	0.95	0.95
0.9	0.95	0.9	0.9	0.95
0.9	0.95	0.9	0.9	0.9
0.9	0.95	0.9	0.9	1.0

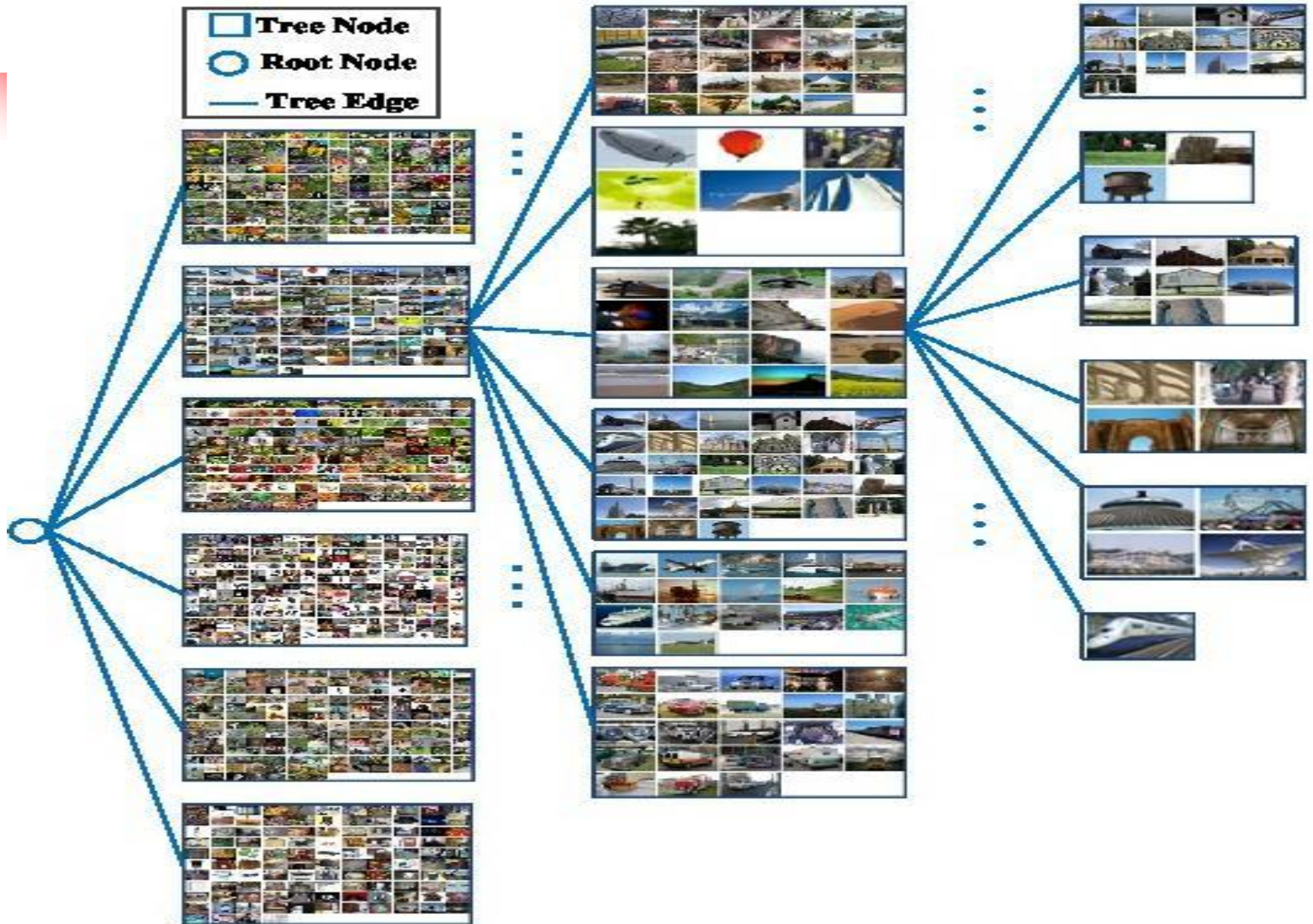
1.0	0.98	0.99
0.98	1.0	0.99
0.99	0.99	1.0

1.0	0.91	0.91	0.9
0.91	1.0	0.91	0.9
0.91	0.91	1.0	0.9
0.9	0.9	0.9	1.0

# 4. Visual Tree Construction: Large-Scale Object Classes



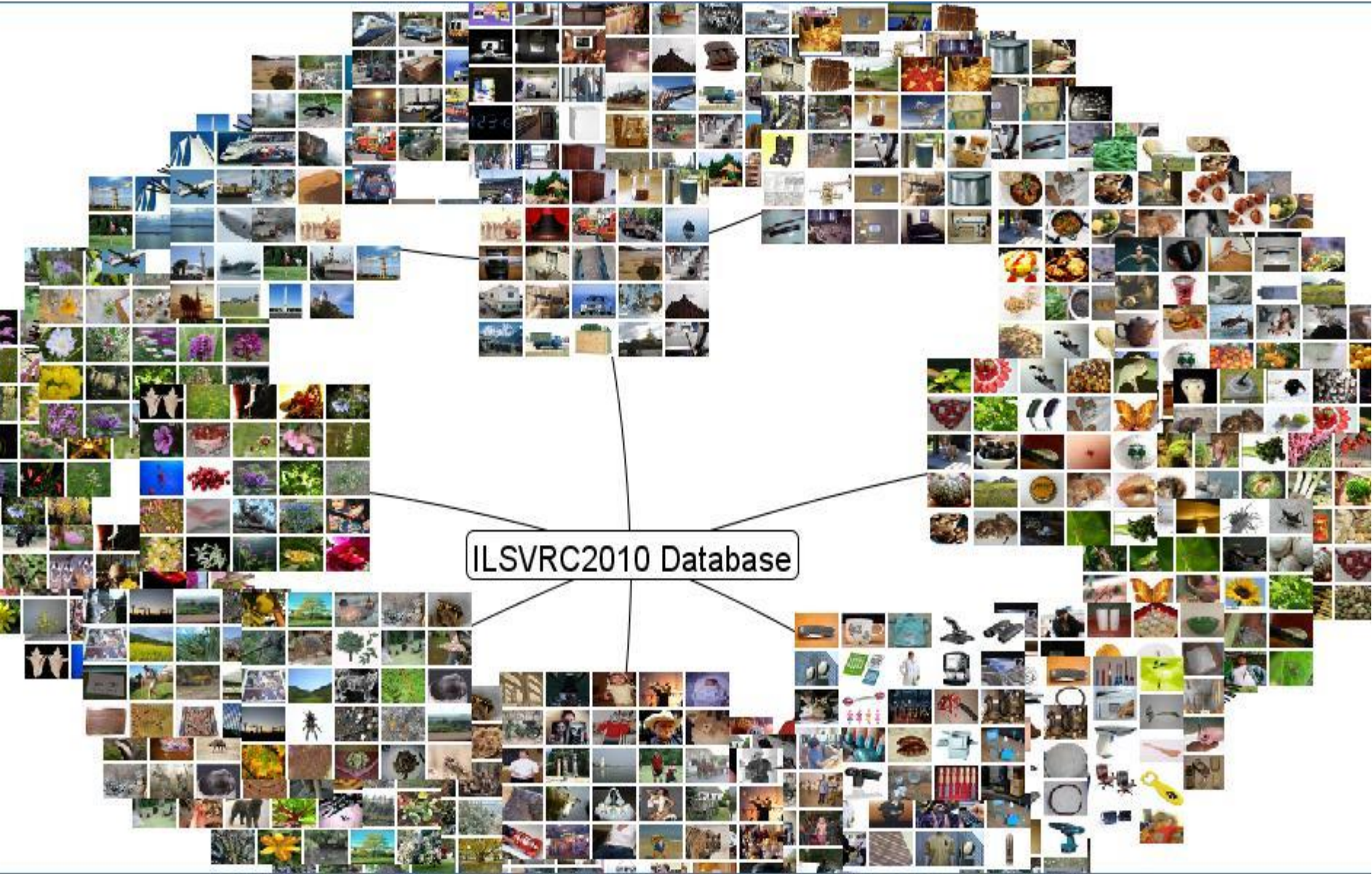
# 4. Visual Tree Construction: Large-Scale Object Classes



# 4. Visual Tree Construction: Large-Scale Object Classes



# 4. Visual Tree Construction: Large-Scale Object Classes





# 4. Visual Tree Construction: Large-Scale Object Classes

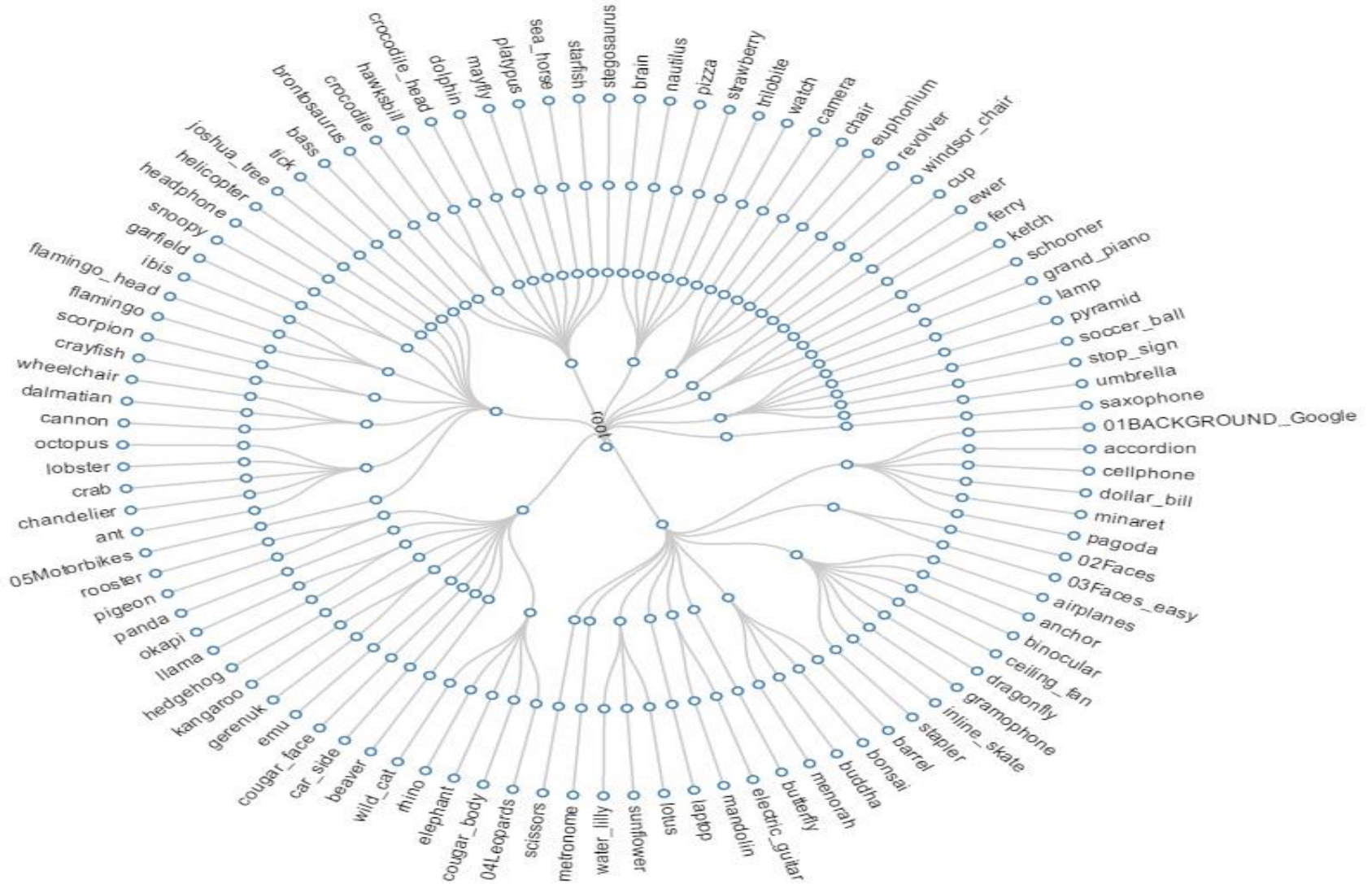


(a)

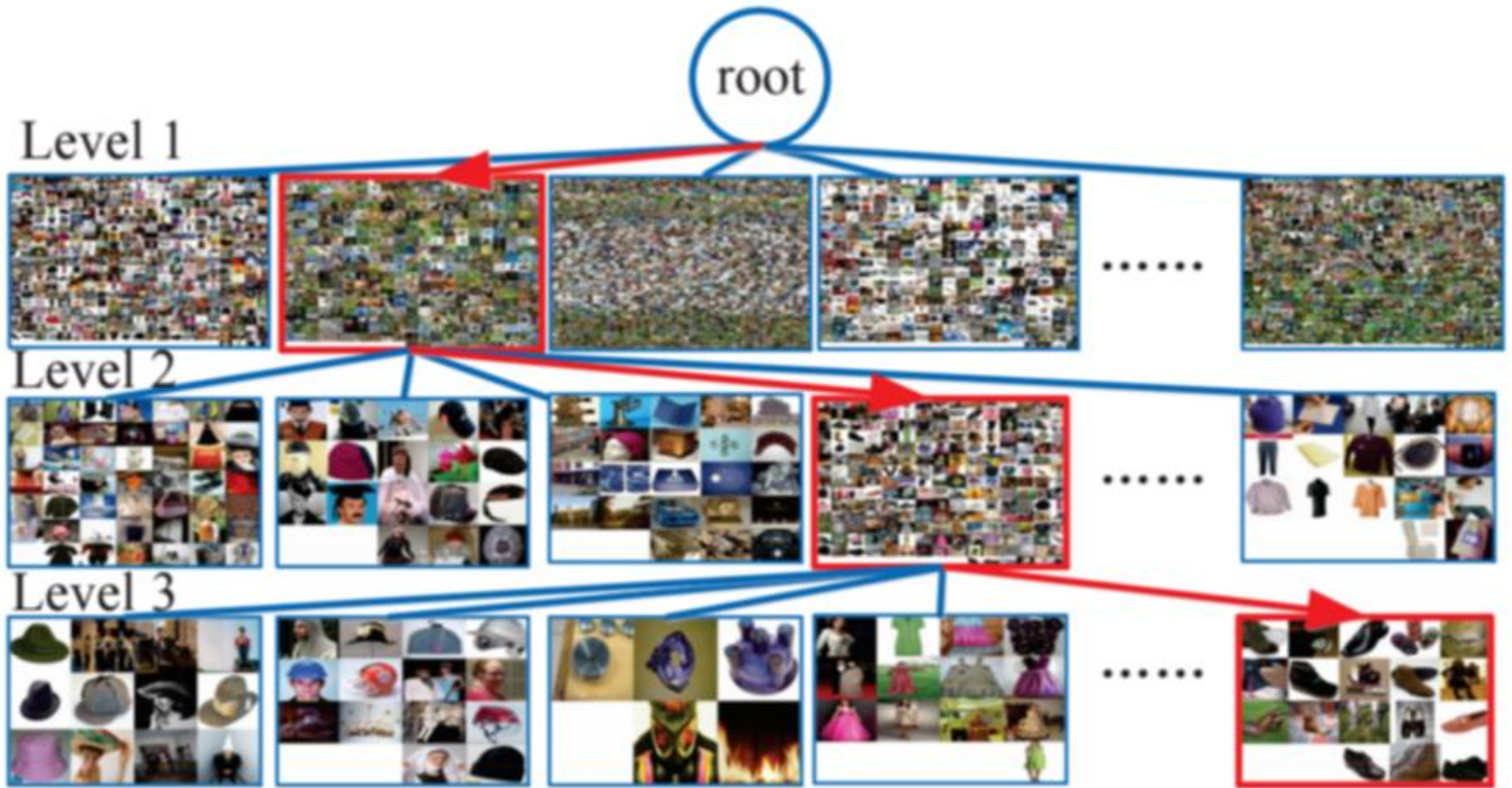


(b)

# 4. Visual Tree for CalTech101



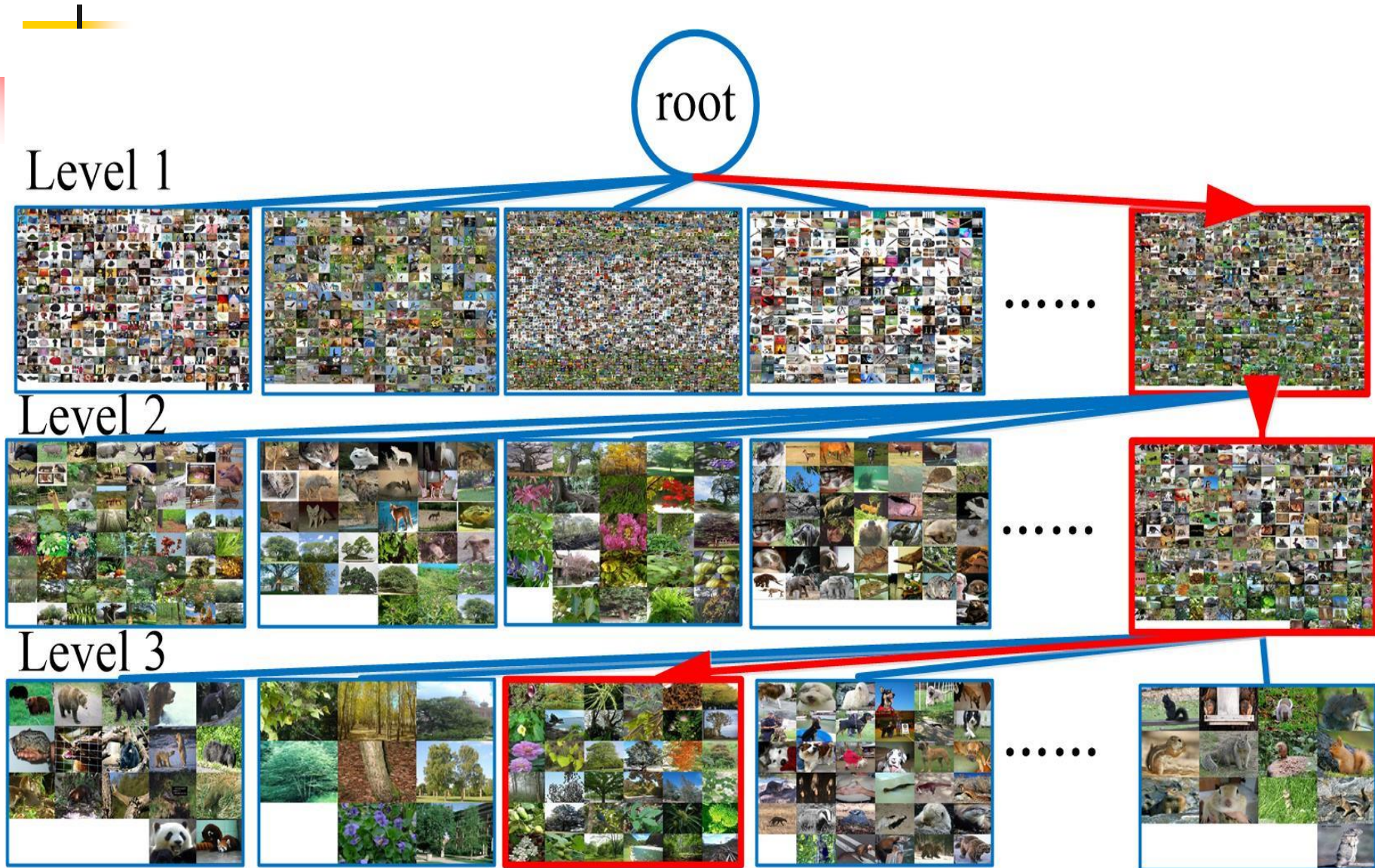
## 4. Visual Tree for ImageNet 10K



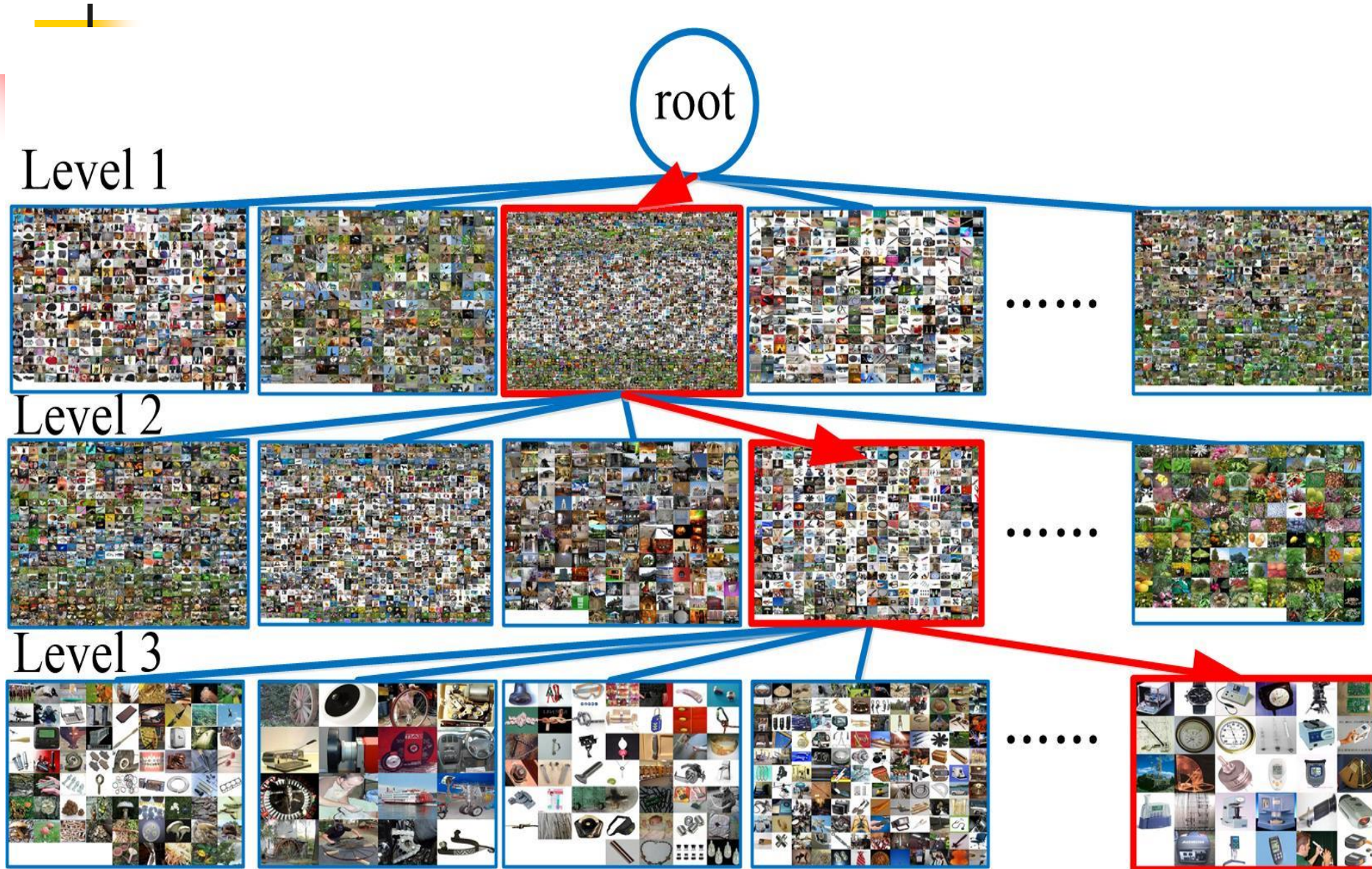
# 4. Visual Tree for ImageNet 10K



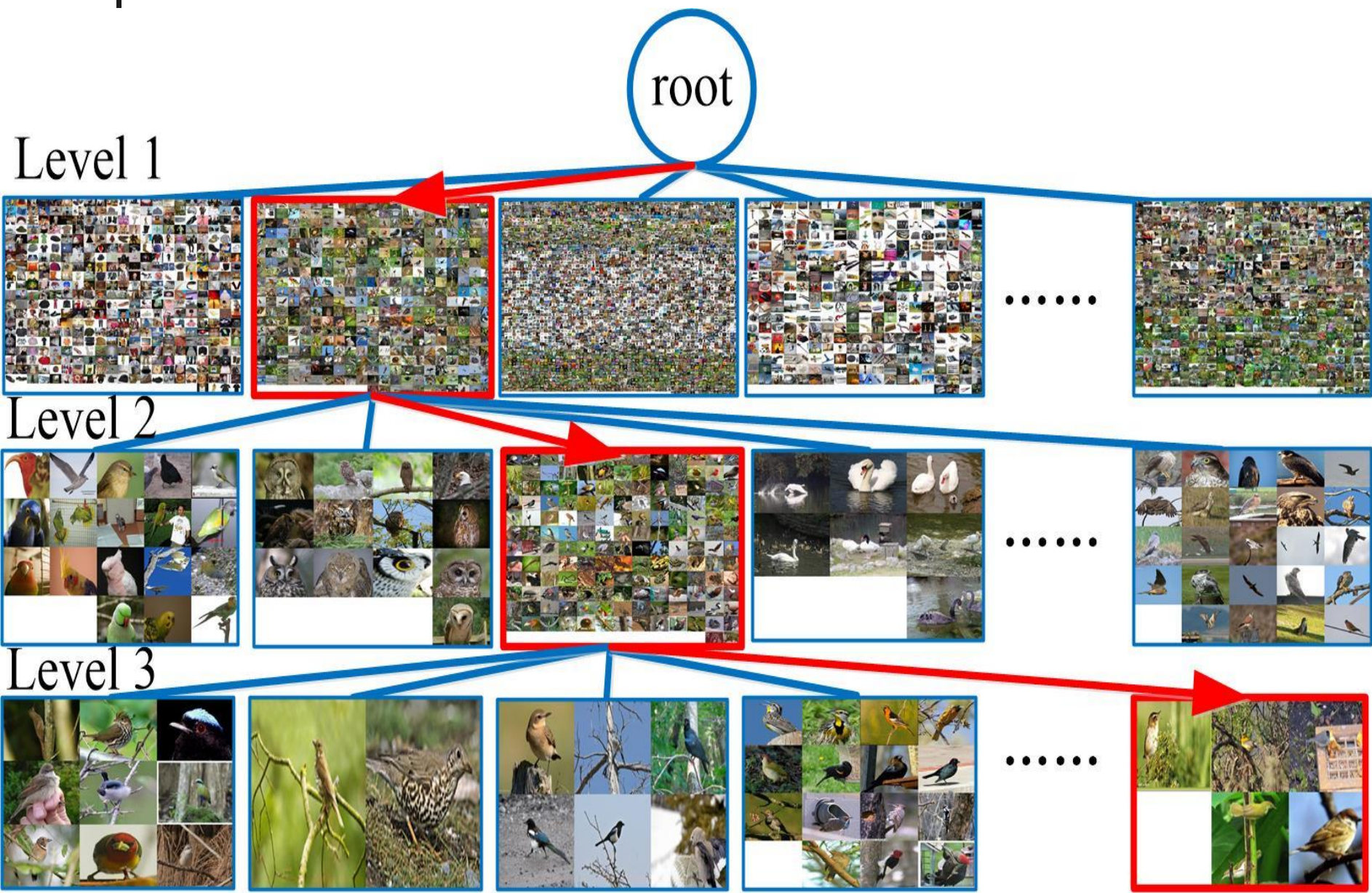
# 4. Visual Tree Construction



# 4. Visual Tree Construction

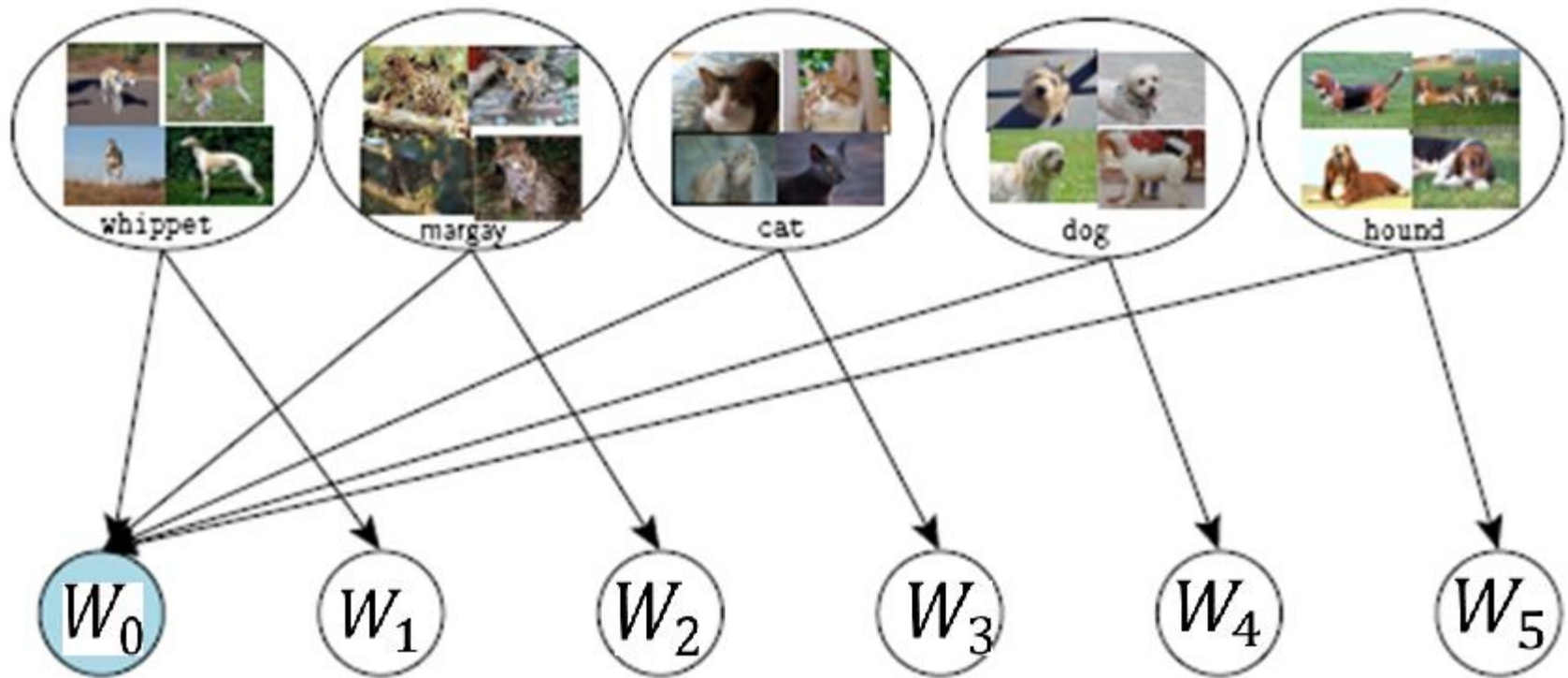


# 4. Visual Tree Construction



# 5. Hierarchical Deep Multi-Task Learning

## ■ Deep Multi-Task Learning







## 5. Hierarchical Deep Multi-Task Learning

---

### ■ Deep Multi-Task Learning

$$\min \left\{ C \sum_{l=1}^R \sum_{j=1}^B \xi_j^l + \delta_1 \text{Tr} (WW^T) + \frac{\delta_2}{2} \text{Tr} (WLW^T) \right\}$$

*subject to:*

$$\forall_{l=1}^R \forall_{j=1}^B : y_j^l (W_j^T \cdot x_j^l + b) \geq 1 - \xi_j^l, \quad \xi_j^l \geq 0$$



## 5. Hierarchical Deep Multi-Task Learning

---

### ■ Deep Multi-Task Learning

$$\min \left\{ \sum_{j=1}^B \sum_{l=1}^R \beta_l^j - \frac{1}{2\delta_1} \beta^T Y \Re \left( \Re + \frac{\delta_2}{\delta_1} \Re (L \otimes I) \Re \right)^{-1} \Re Y \beta \right\}$$

*subject to:*

$$\forall_{l=1}^R \forall_{j=1}^B : \sum_{l=1}^R \beta_l^j \cdot y_l^j = 0, \quad 0 \leq \beta_l^j \leq 1$$



## 5. Hierarchical Deep Multi-Task Learning

---

### ■ Deep Multi-Task Learning

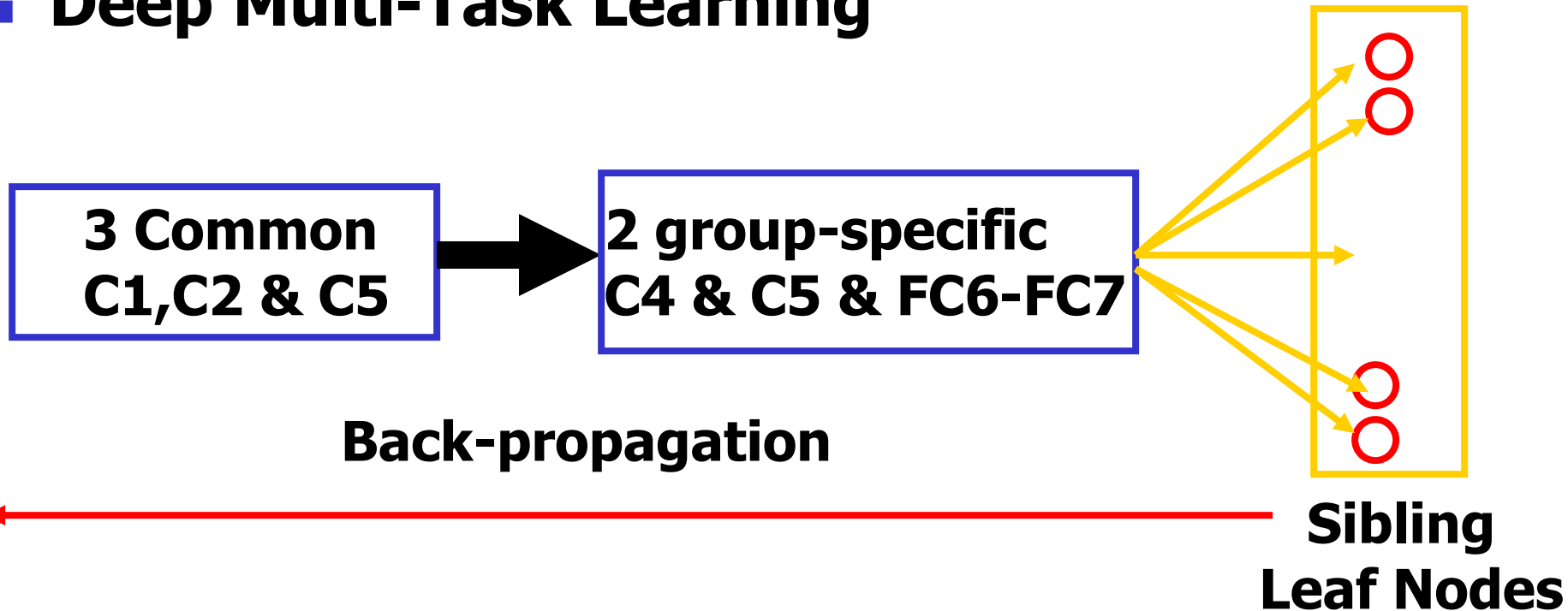
$$\alpha^* = \frac{1}{2\delta_1} \left( \mathfrak{R} + \frac{\delta_2}{\delta_1} \left( \mathfrak{R} \left( L \otimes I \right) \mathfrak{R} \right)^{-1} \mathfrak{R} Y \beta^* \right)$$

### Multi-Task Classifiers at Sibling Leaf Nodes

$$\forall_{j=1}^B : f_{c_j}^1(x) |_{F_{c_j}^1} = \sum_{l=1}^R \alpha_j^{l*} \kappa(x_j^l, x) + b_j^*, \quad c_j \in c_h$$

# 5. Hierarchical Deep Multi-Task Learning

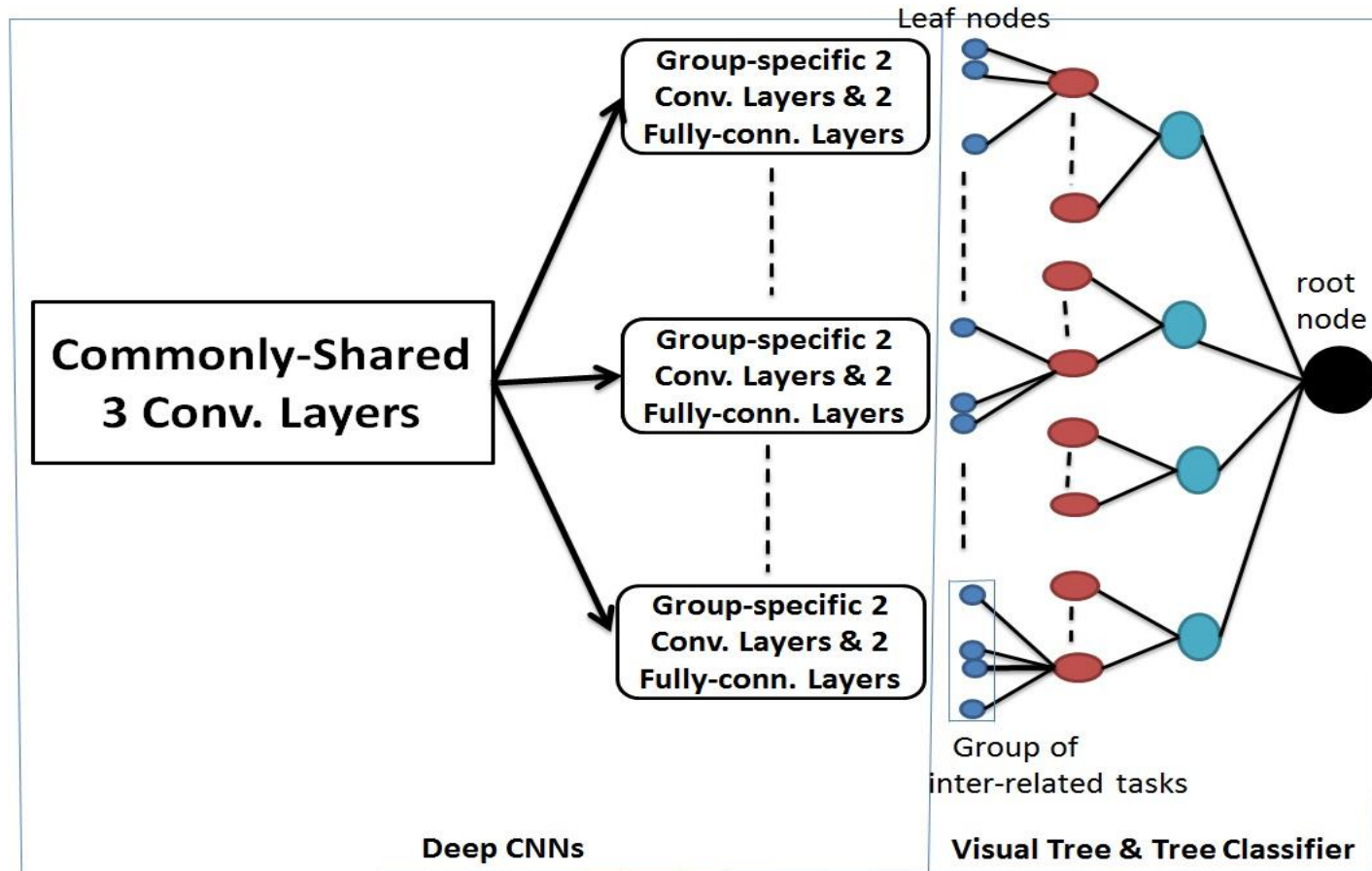
## ■ Deep Multi-Task Learning



$$\frac{\partial \mathcal{L}(W, X, Y)}{\partial W} = W(\delta_1 I + \delta_2 L) - \mu \sum_{l=1}^R \sum_{j=1}^B I\{y_j^l\} \left[ x_j^l - \sum_{j=1}^B x_j^l \right]$$

# 5. Hierarchical Deep Multi-Task Learning

## ■ Hierarchical Deep Multi-Task Learning



## 5. Hierarchical Deep Multi-Task Learning

### ■ Hierarchical Deep Multi-Task Learning

$$\min \left\{ C \sum_{m=1}^R \sum_{h=1}^B \xi_j^m + \gamma_1 \text{Tr} (W W^T) + \frac{\gamma_2}{2} \text{Tr} (W L W^T) \right\}$$

*subject to:*

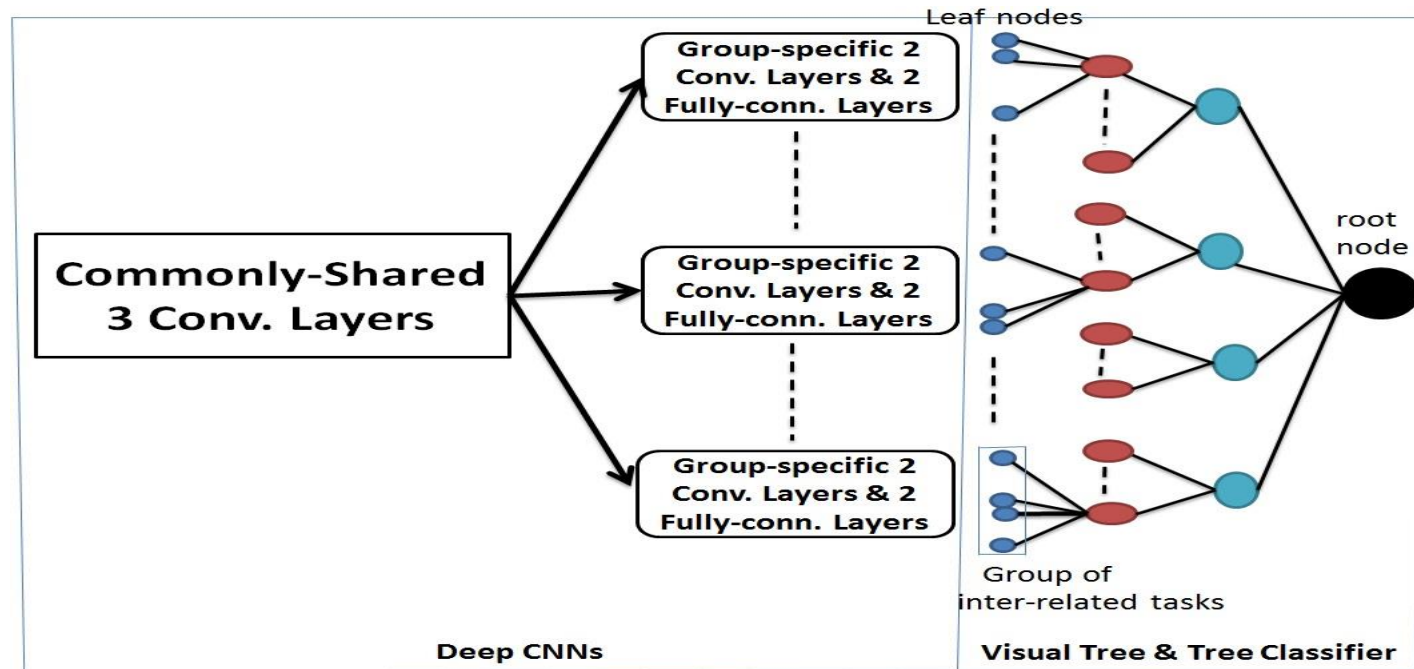
$$\forall_{m=1}^R \forall_{h=1}^B : y_h^m (W_h^T \cdot x_h^m + b) \geq 1 - \xi_h^m, \xi_h^m \geq 0, c_h \in c_k$$

$$\forall_{h=1}^B : f_{c_h}^{l+1}(x) |_{F_{c_h}^{l+1}} - f_{c_j}^l(x) |_{F_{c_j}^l} \geq 0$$

$$\forall_{h=1}^B : f_{c_h}^{l+1}(x) |_{F_{c_h}^{l+1}} = \sum_{j=1}^B \eta_j f_{c_j}^l(x) |_{F_{c_j}^l}$$

# 5. Hierarchical Deep Multi-Task Learning

## ■ Hierarchical Deep Multi-Task Learning



$$\frac{\partial \Psi(W, X, Y)}{\partial W} = \Omega(\gamma_1 I + \gamma_2 \Xi) - \nu \sum_{m=1}^N \sum_{h=1}^L I\{y_h^m\} \left[ x_h^m - \sum_{h=1}^L x_h^m \right] + \lambda$$



## 5. Hierarchical Deep Multi-Task Learning

---

- **Back Propagation**

- **Errors from High-Level Node**

- **Node classifier for itself**
- **Node Classifiers for lower-level nodes which treat it as their ancestors**
- **Weights of deep networks**

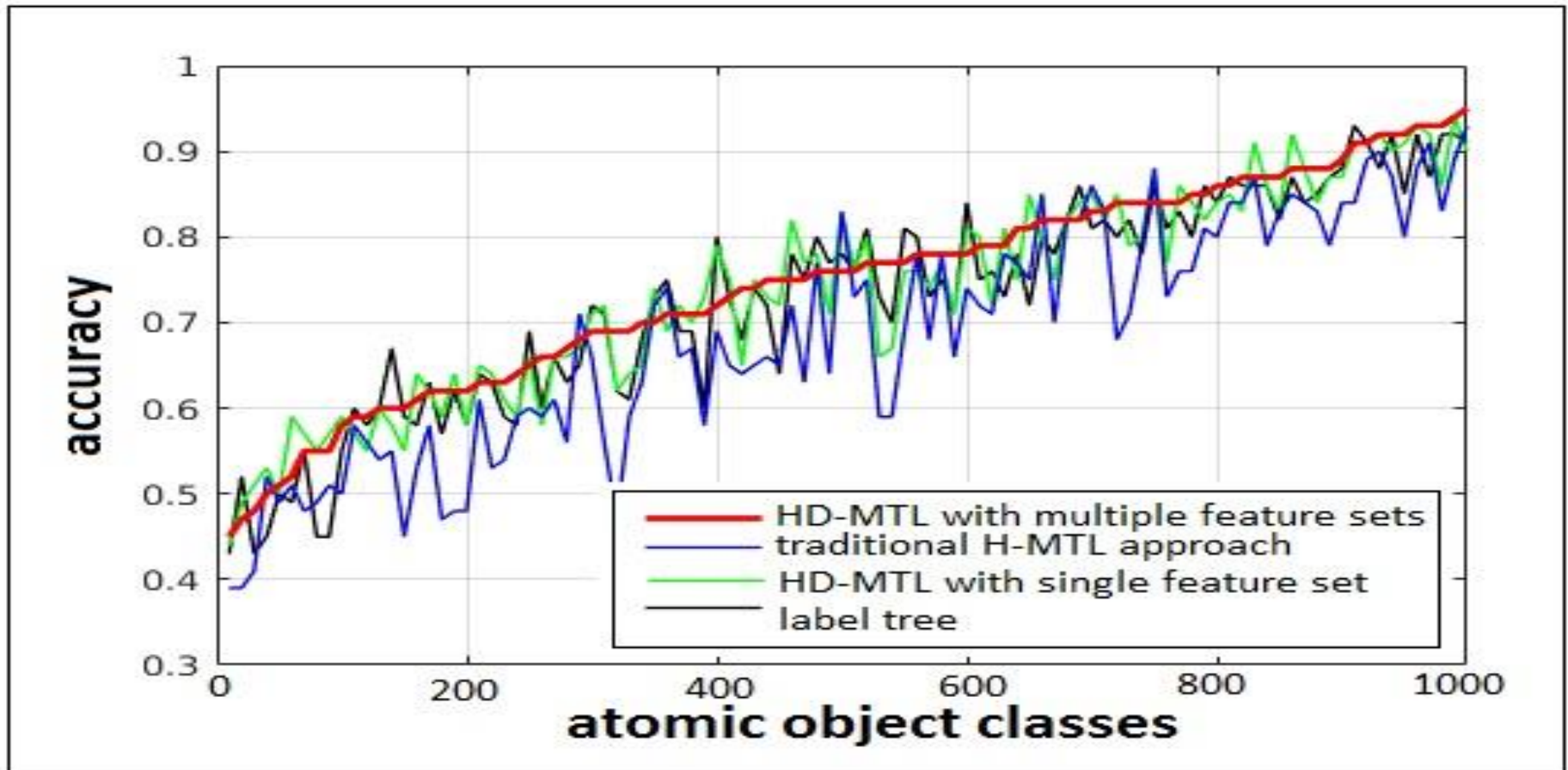
- **Errors from Leaf Node**

- **Node classifier for itself**
- **Weights of deep networks**



## 6. Some Experimental Results

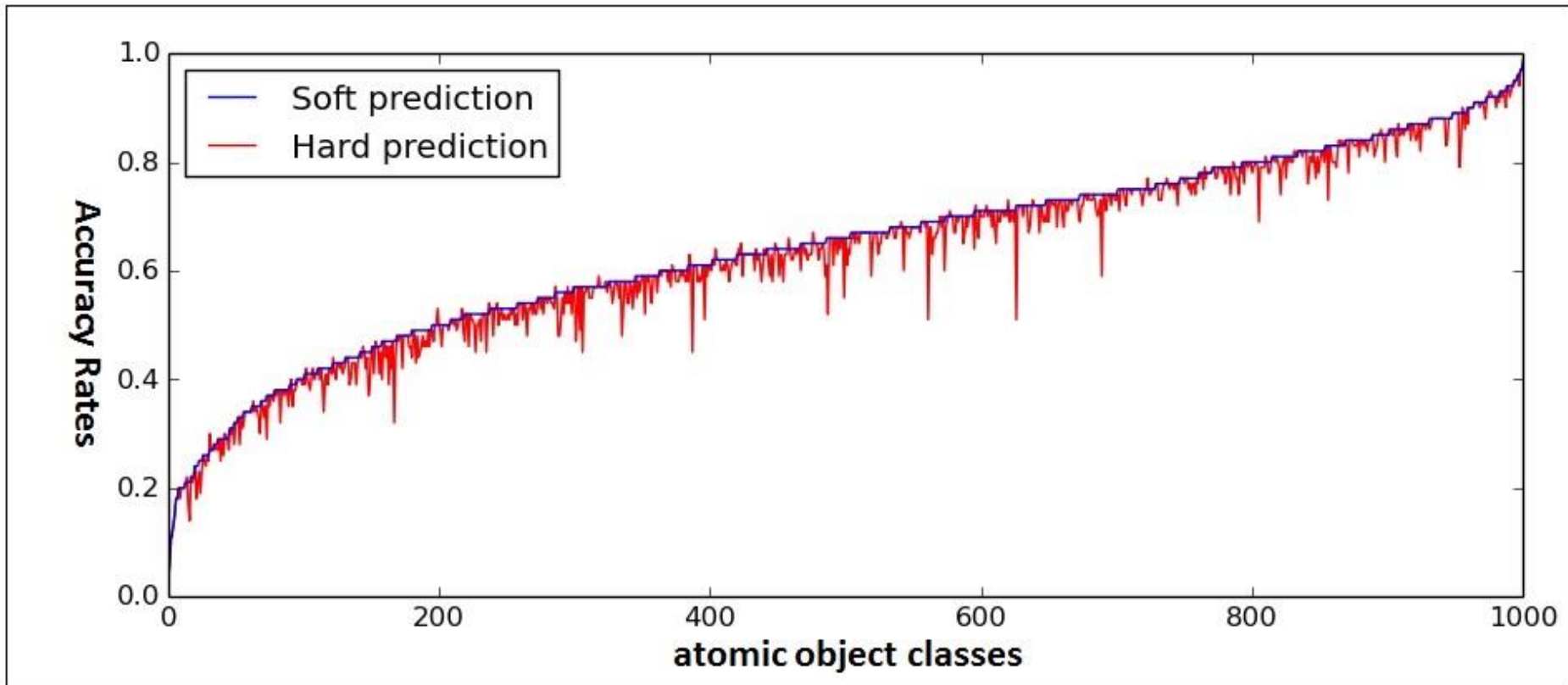
### ■ Impacts of Feature Subset Selection





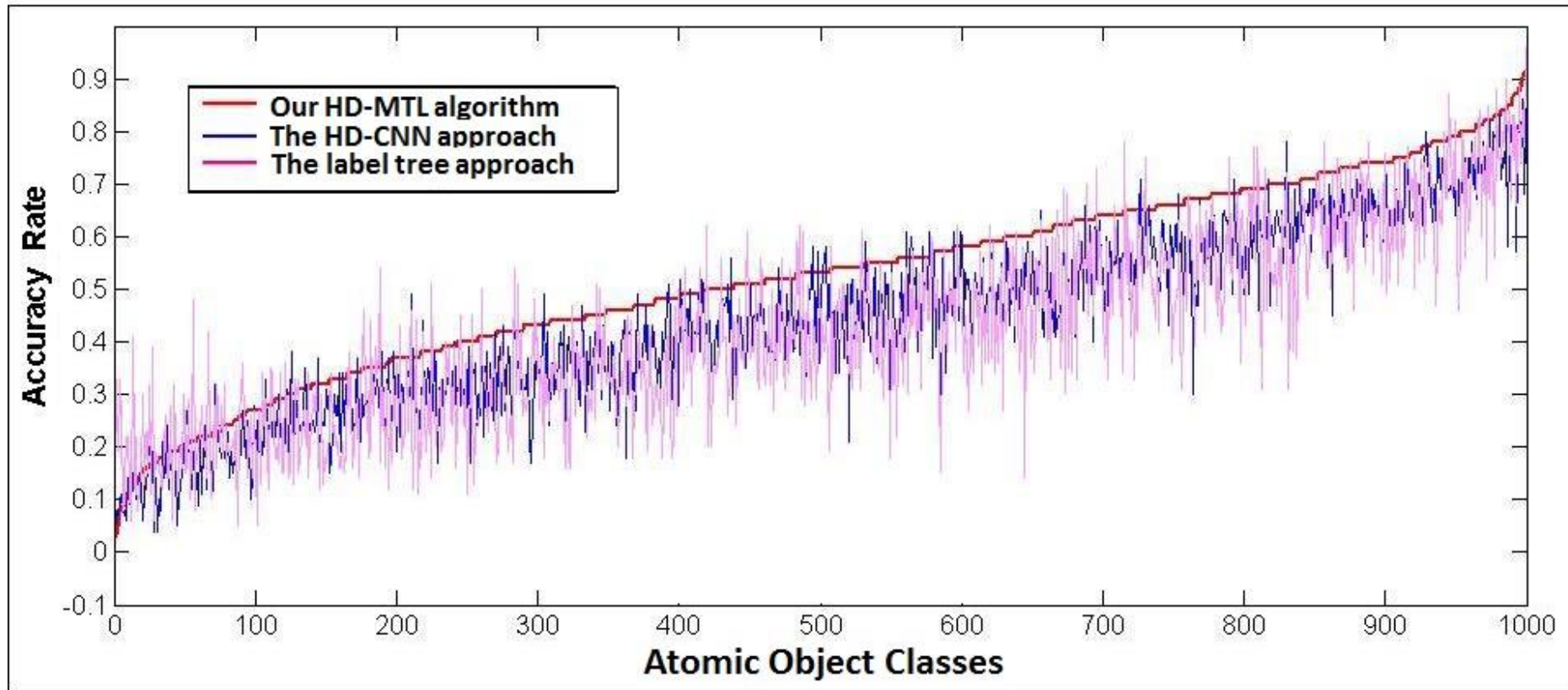
## 6. Some Experimental Results

### ■ Impacts of Soft Prediction



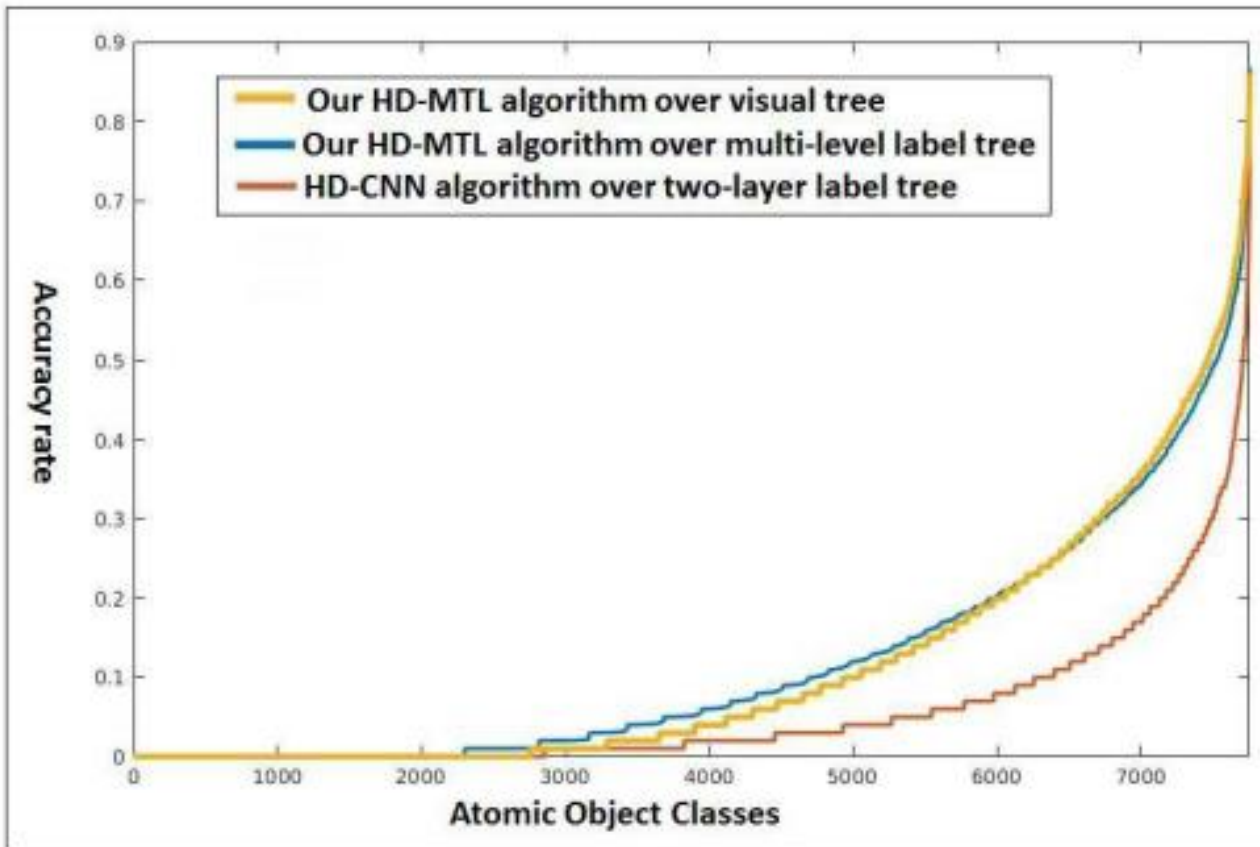
## 6. Some Experimental Results

### ■ Impacts of Deep Multi-Task Learning



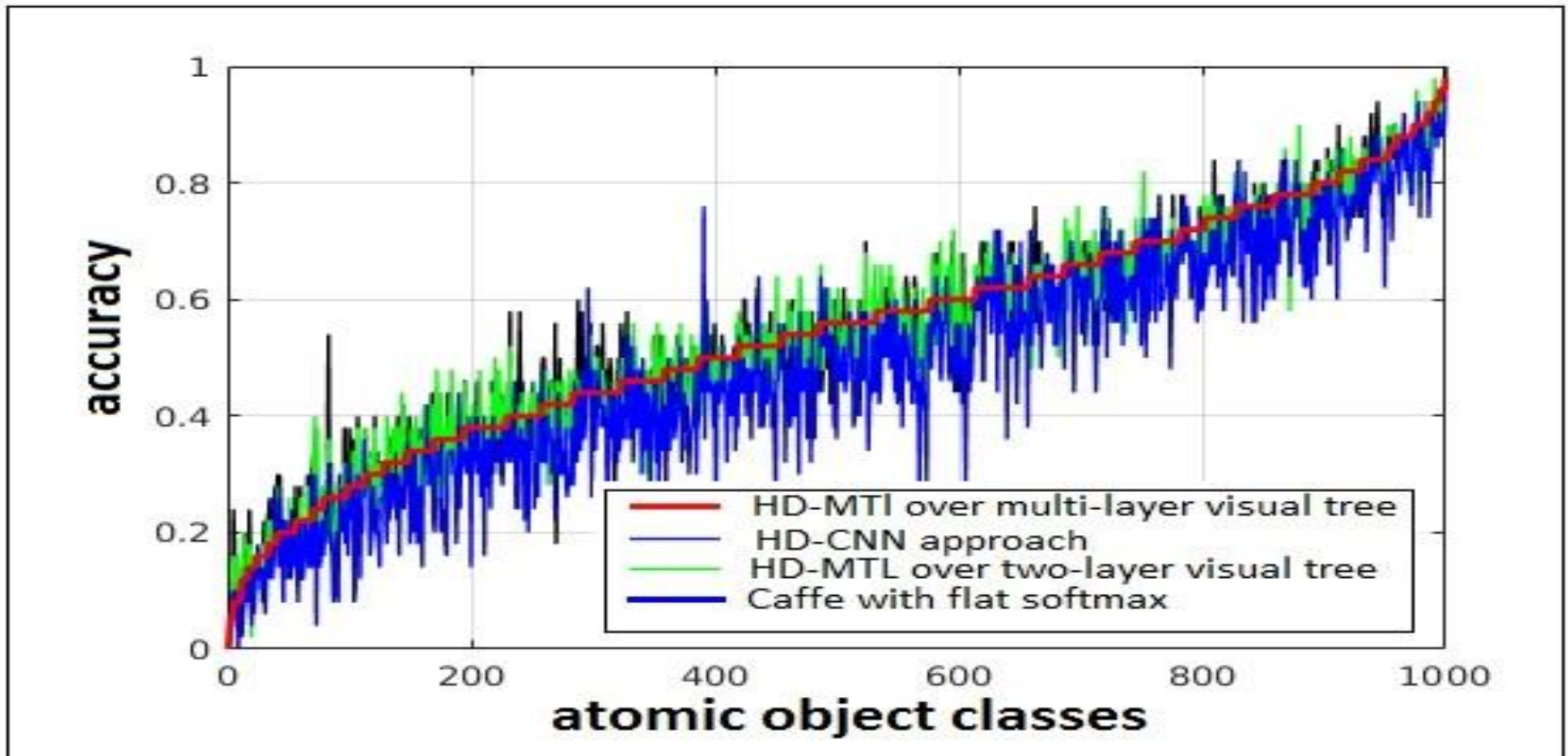
## 6. Some Experimental Results

### ■ Impacts of Deep Multi-Task Learning



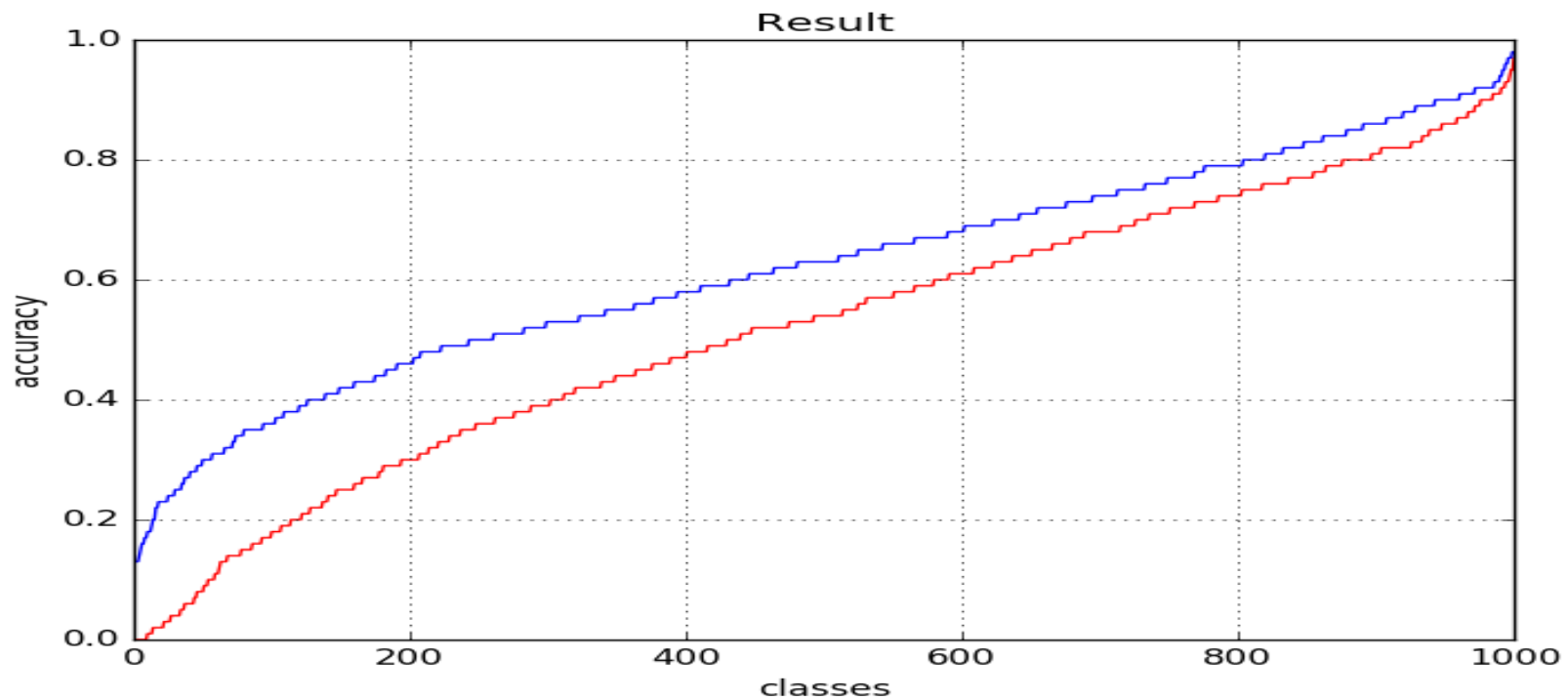
## 6. Some Experimental Results

### ■ Impacts of Visual Tree



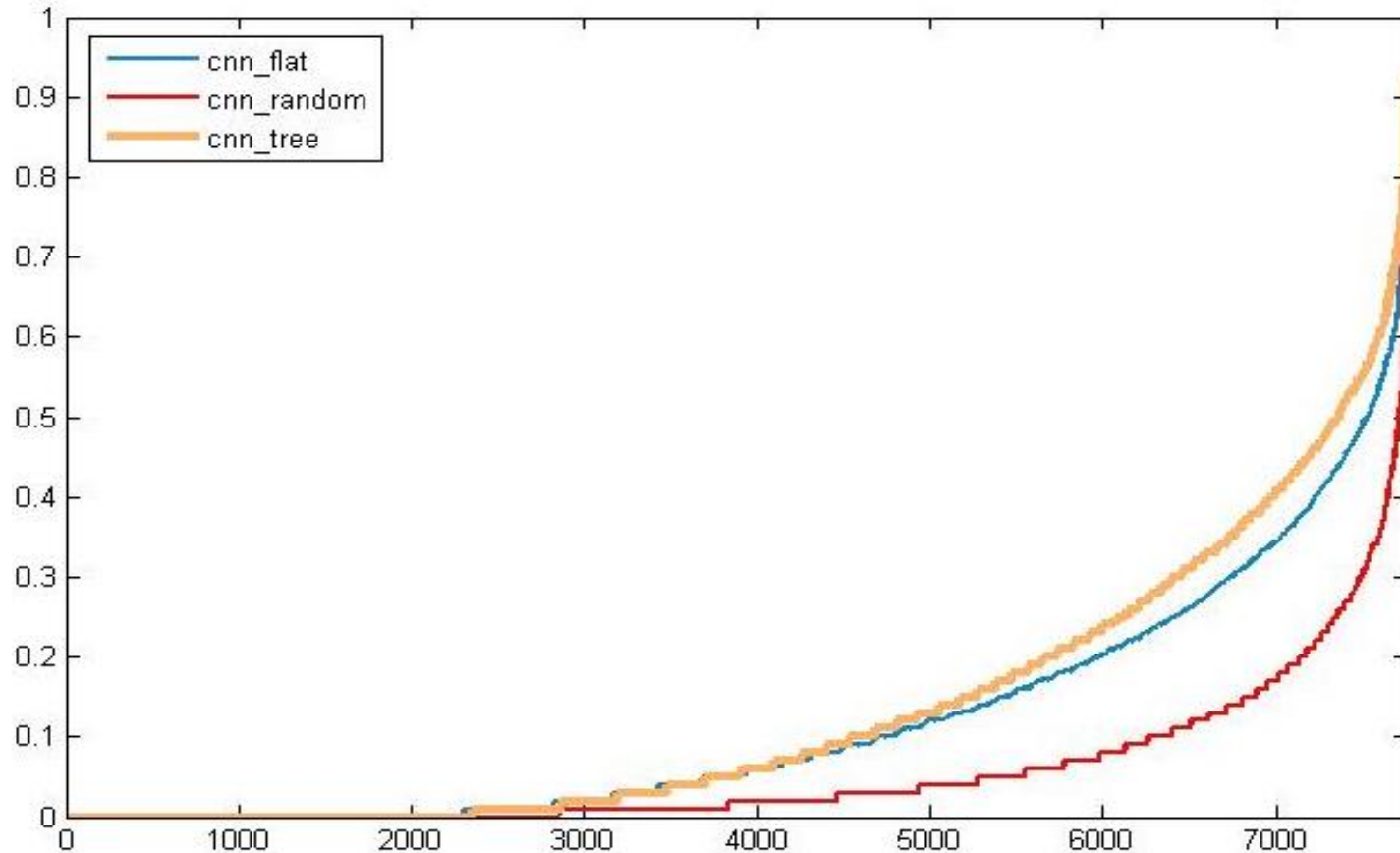
## 6. Some Experimental Results

### ■ Impacts of Visual Tree



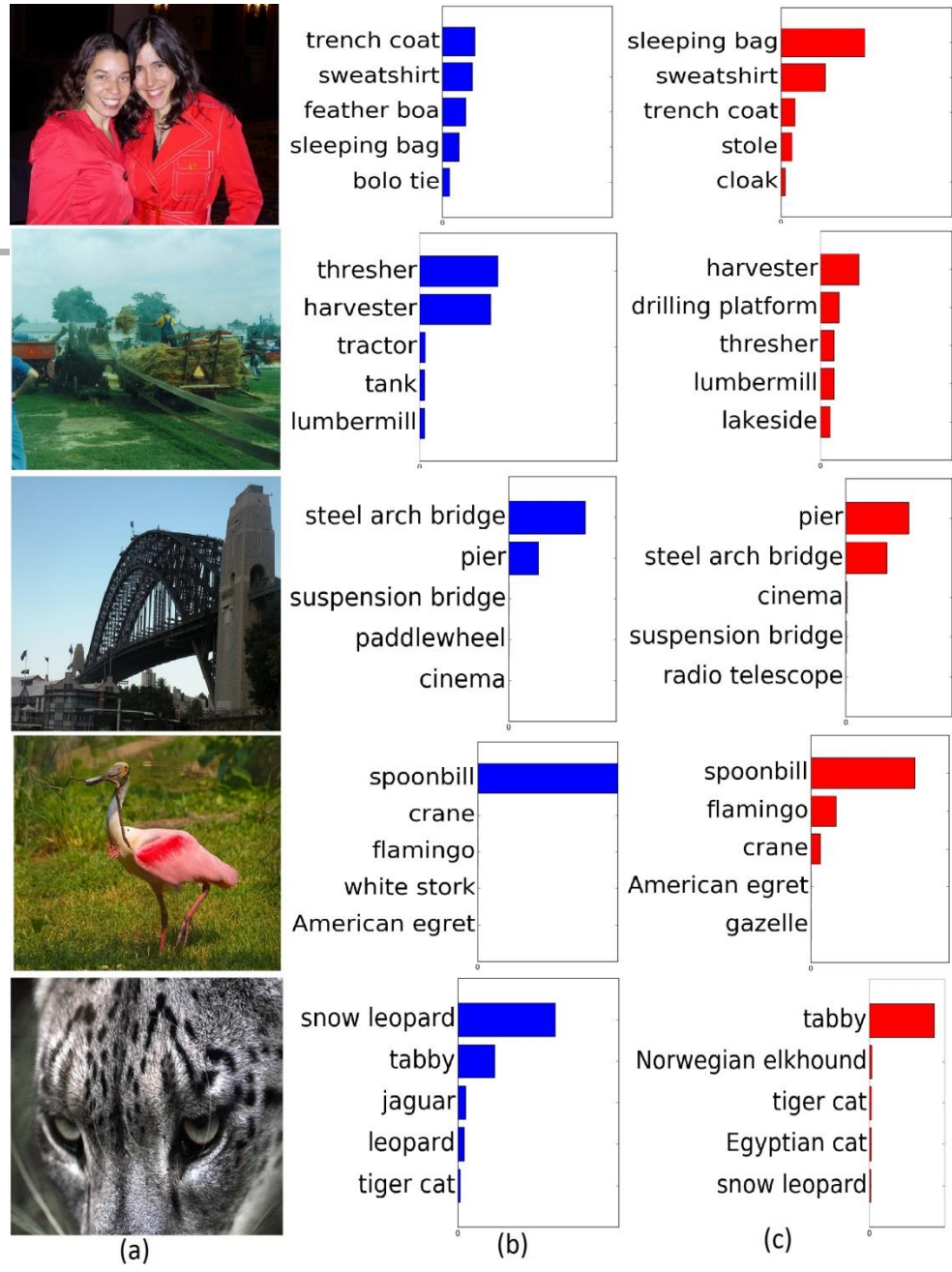
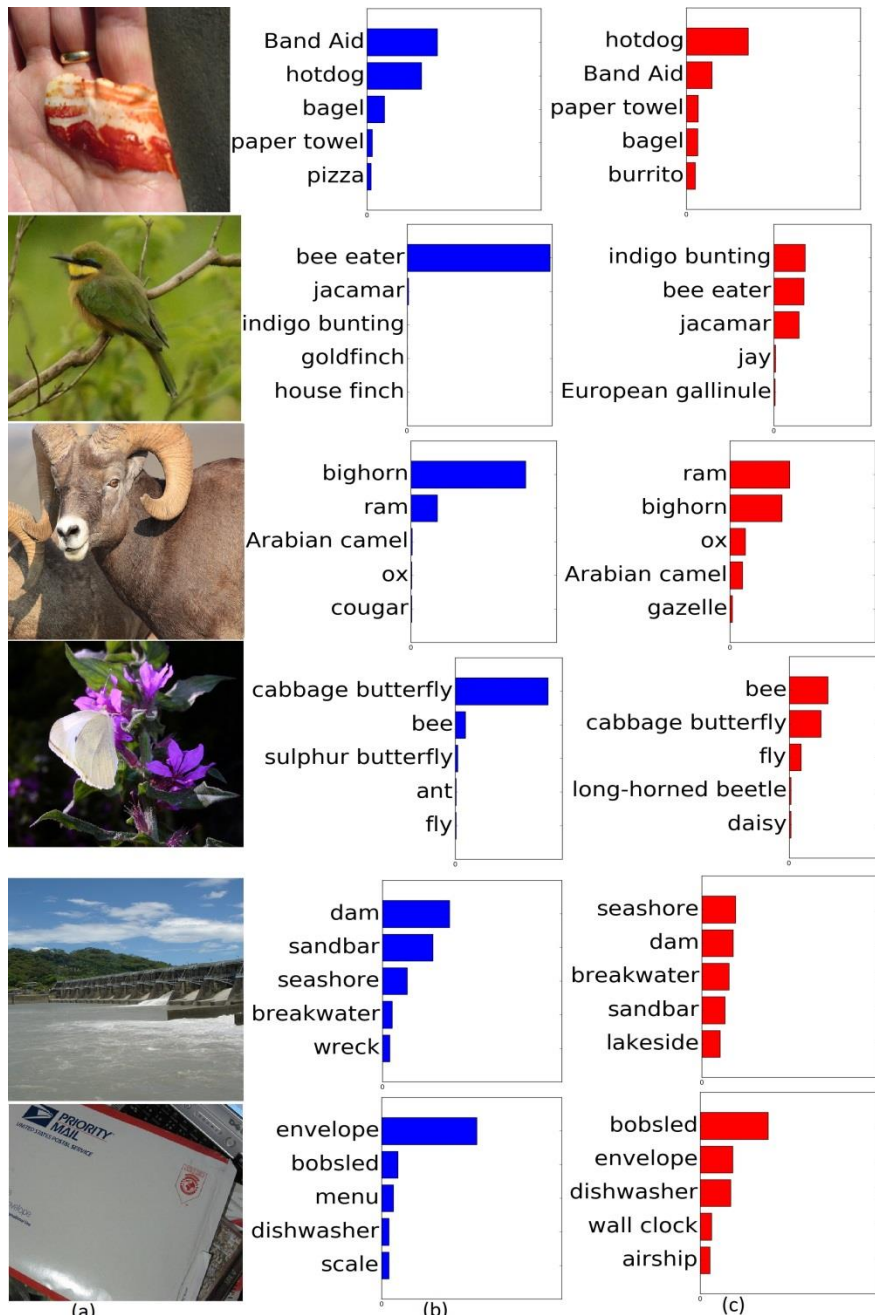
## 6. Some Experimental Results

### ■ Impacts of Visual Tree





# Prediction Confidence Enhancement



(a)

(b)

(c)

(a)

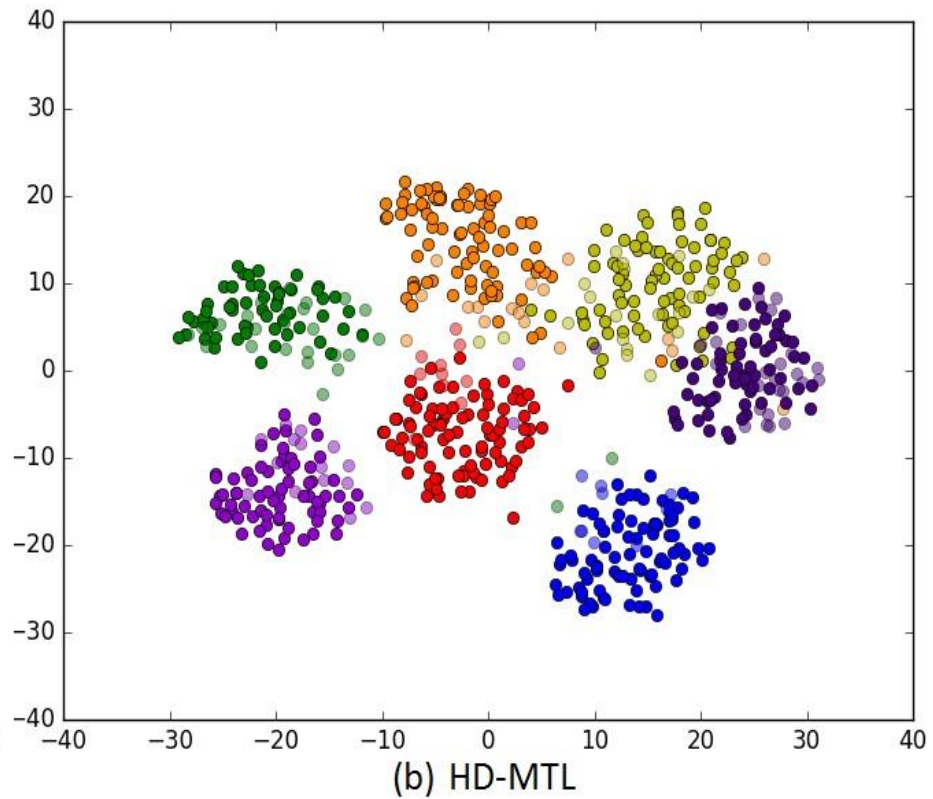
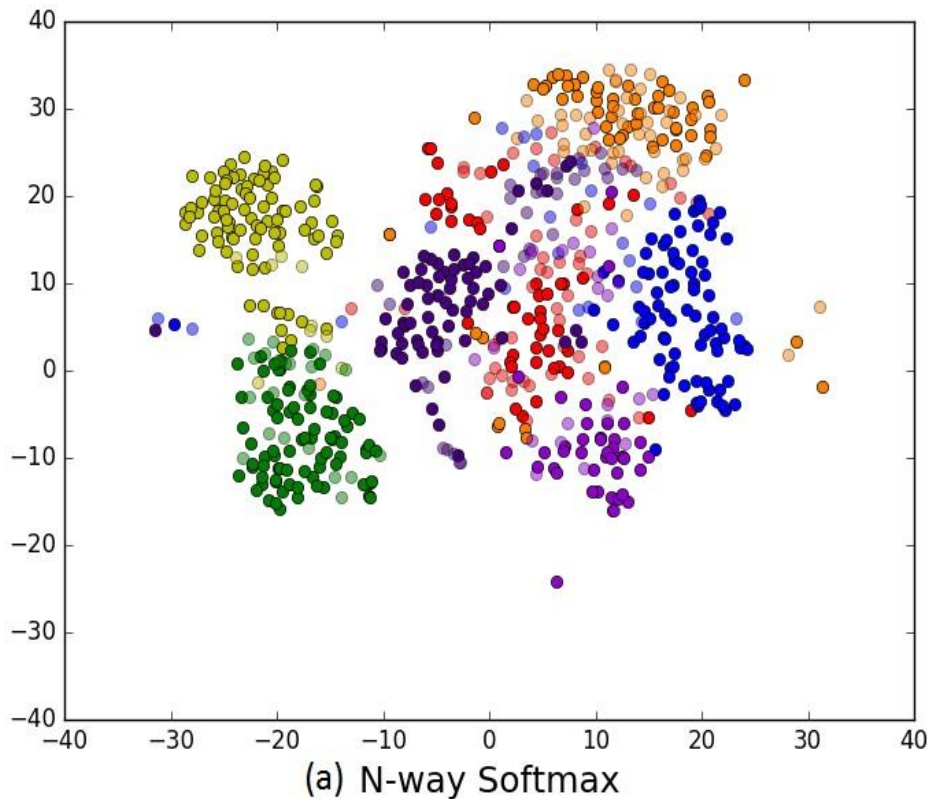
(b)

(c)



## 6. Some Experimental Results

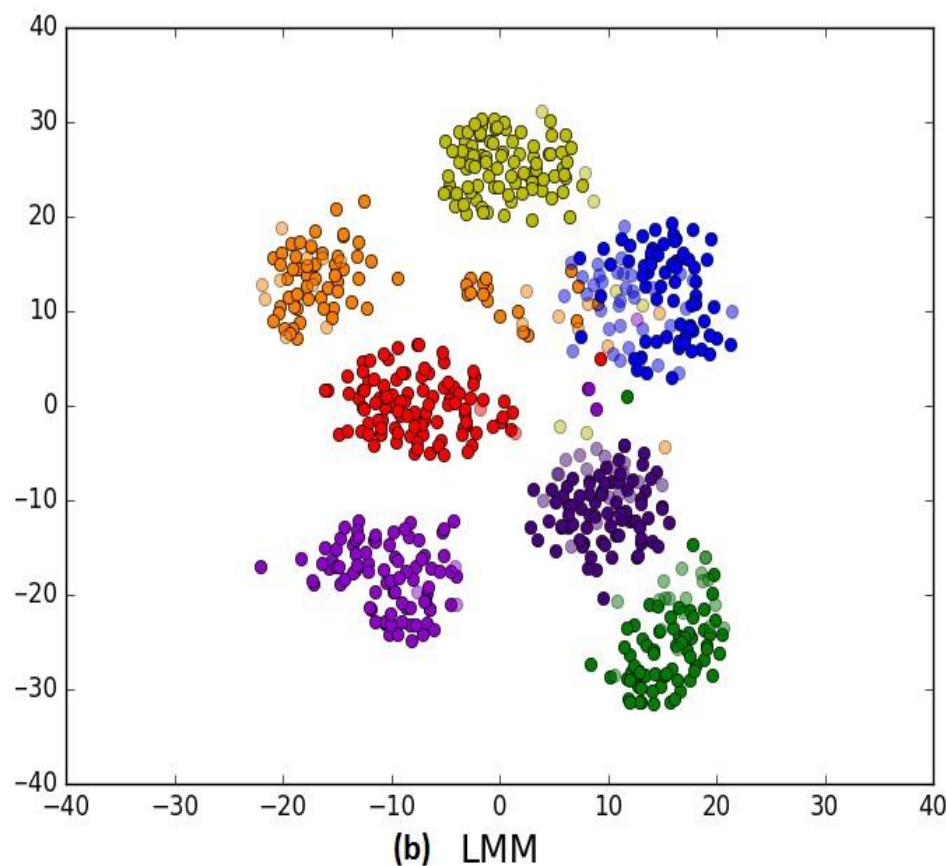
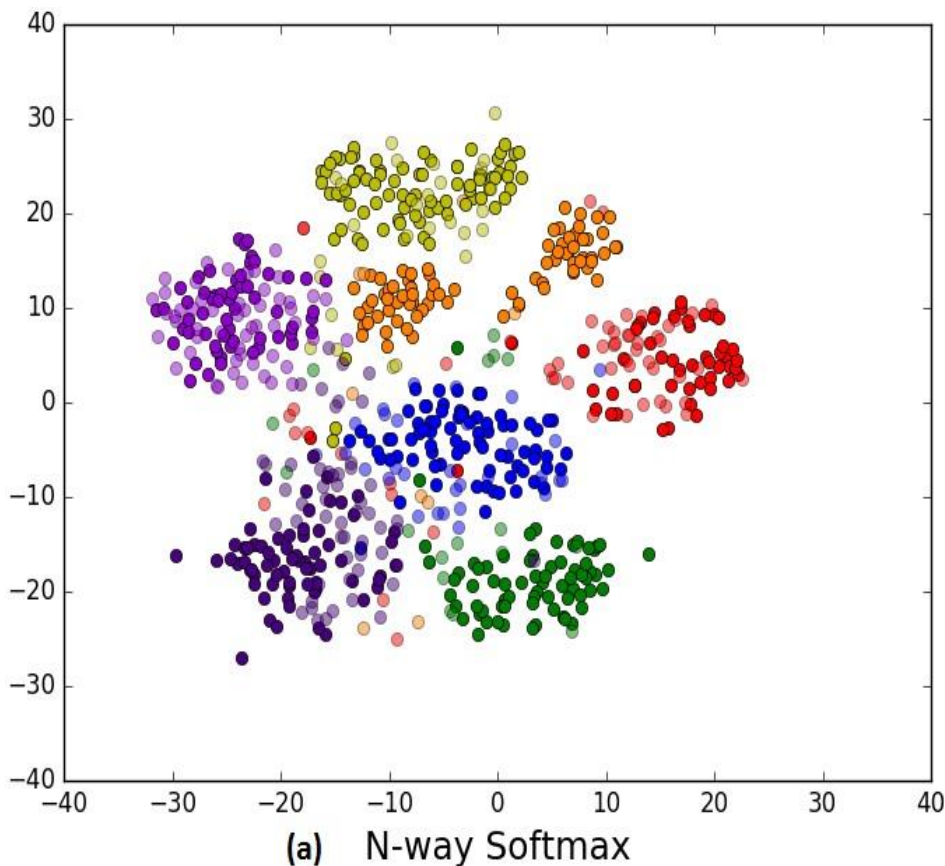
### ■ Impacts of Deep Multi-Task Learning





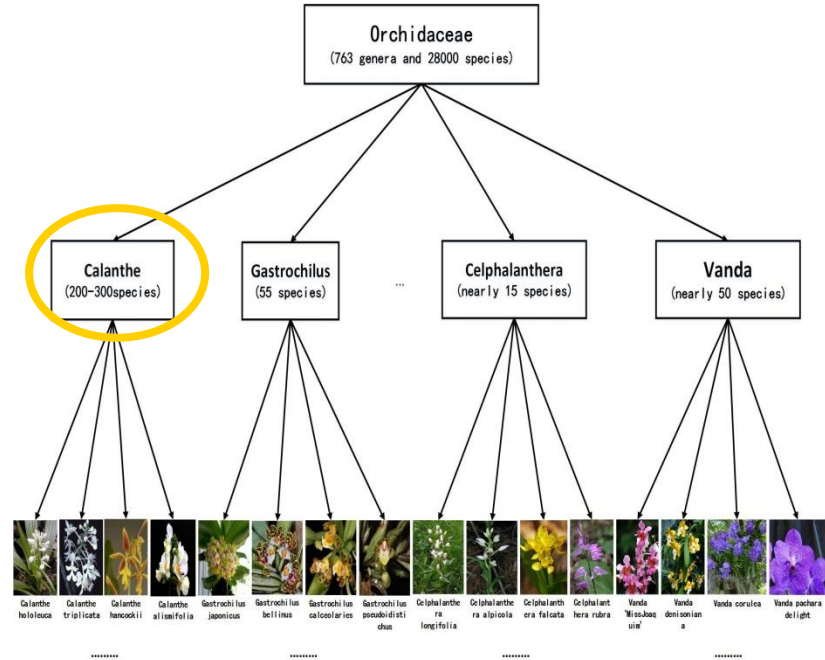
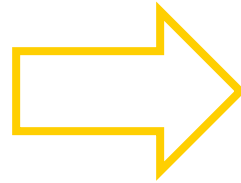
## 6. Some Experimental Results

### ■ Impacts of Deep Multi-Task Learning



# Concept Ontology vs. Visual Tree

- **Early Stop**



- **Semantic Interpretation**