

Improving Multi-label Learning with Missing Labels by Structured Semantic Correlations

Hao Yang¹, Joey Tianyi Zhou² and Jianfei Cai¹

¹NTU, ²Sony Research

Multi-label Image Recognition IS Everywhere

- All images can be assigned with multiple labels/tags.
 - Better understanding for the images



©jaymiheimbuch.com

Dog ?



©jaymiheimbuch.com



©jaymiheimbuch.com

Dog, chair
carpet, indoor,
high five, pet,
funny...

Multi-label Image Recognition IS Everywhere

- All images can be assigned with multiple labels/tags.
 - Better understanding for the images
 - More convenient for possible NLP related applications



A baby eating a piece of paper?

Image Captioning & Visual
Question Answering

Multi-label Image Recognition IS Everywhere

- All images can be assigned with multiple labels/tags.
 - Better understanding for the images
 - More convenient for possible NLP related applications
 - Easier to retrieve relevant images

Explore / Tags / airshow

Sort by:
Most recent • **Most interesting**

airshow clusters

Explore and refine this airshow list with our wonderful cluster goodness!

Related tags:

aircraft airplane jet plane aviation fighter f16 military sky



From schaeffers...



From schaeffers...



From obermiller...



From schaeffers...

Your photostream / Tags / airshow

See all public content tagged with airshow

Edit these in a batch

Change this tag?

What are tags?

You can give your uploads a "tag", which is like a keyword. Tags help you find things which have something in common. You can assign up to 75 tags to each photo or video.



Want to format your comment?

PREVIEW

POST COMMENT

EasyShare C743 Zoom. (add to map)

11 views

This photo belongs to

markembling's photostream (33)

This photo also appears in

Farnborough Air Show (set: 37)



Advertisement

Get Sorted with Organizr

Batch edit groups of photos and videos

Batch Organizr

Drag items here to add them all in batch. You can then change any attributes or create a new set.

All your photos



flickr

Want to go ad-free? Get Flickr Pro!

People in this photo (add a person)

Adding people shares who is in your photo

Tags (add a tag)

air-show • farnborough • red-arrows

Multi-label learning

- Provides many possibilities
- Poses many challenges

Challenges

- Feature Extraction
 - Both global and local level features are important
- Label related challenges
 - Large label space
 - Noisy labels
 - Missing labels

Challenges

- Feature Extraction
 - Both global and local level features are important
- Label related challenges
 - Large label space
 - Noisy labels
 - Missing labels

Missing Labels Problem

- Inevitable for multi-label image recognition
 - The number of possible labels/tags could be large.
 - There often exists ambiguity among labels

Missing Labels Problem

- Inevitable for multi-label image recognition
 - The number of possible labels/tags could be large.
 - There often exists ambiguity among labels



Missing Labels Problem

- Inevitable for multi-label image recognition
 - The number of possible labels/tags could be large.
 - There often exists ambiguity among labels



**animal, clouds, plantlife,
sky, water**

Missing Labels Problem

- Inevitable for multi-label image recognition
 - The number of possible labels/tags could be large.
 - There often exists ambiguity among labels



Missing Labels Problem

- Inevitable for multi-label image recognition
 - The number of possible labels/tags could be large.
 - There often exists ambiguity among labels



*grass, green, lake,
landscape, reindeer*

Missing Labels Problem

- Naturally many works have been proposed to deal with the problem.
 - Instance – label correlations
 - Label – label correlations
 - Instance – instance correlations

Missing Labels Problem

- Naturally many works have been proposed to deal with the problem.

- Instance – label correlations

- Label – label correlations
- Instance – instance correlations

Missing Labels Problem

- Naturally many works have been proposed to deal with the problem.
 - Instance – label correlations

- Label – label correlations
- Instance – instance correlations



Missing Labels Problem

- Naturally many works have been proposed to deal with the problem.
 - Instance – label correlations

- Label – label correlations
- Instance – instance correlations



Exploit deeper knowledge

Important for missing labels problem

Missing Labels Problem

- Naturally many works have been proposed to deal with the problem.
 - Instance – label correlations

- Label – label correlations
- Instance – instance correlations



Exploit deeper knowledge

Important for missing labels problem

Most existing works

- Only consider linear correlations
- Inefficient for large scale label matrix

Missing Labels Problem

- Naturally many works have been proposed to deal with the problem.
 - Instance – label correlations

- Label – label correlations
- Instance – instance correlations



Exploit deeper knowledge

Important for missing labels problem

Missing Labels Problem

- Naturally many works have been proposed to deal with the problem.
 - Instance – label correlations

- Label – label correlations
- Instance – instance correlations

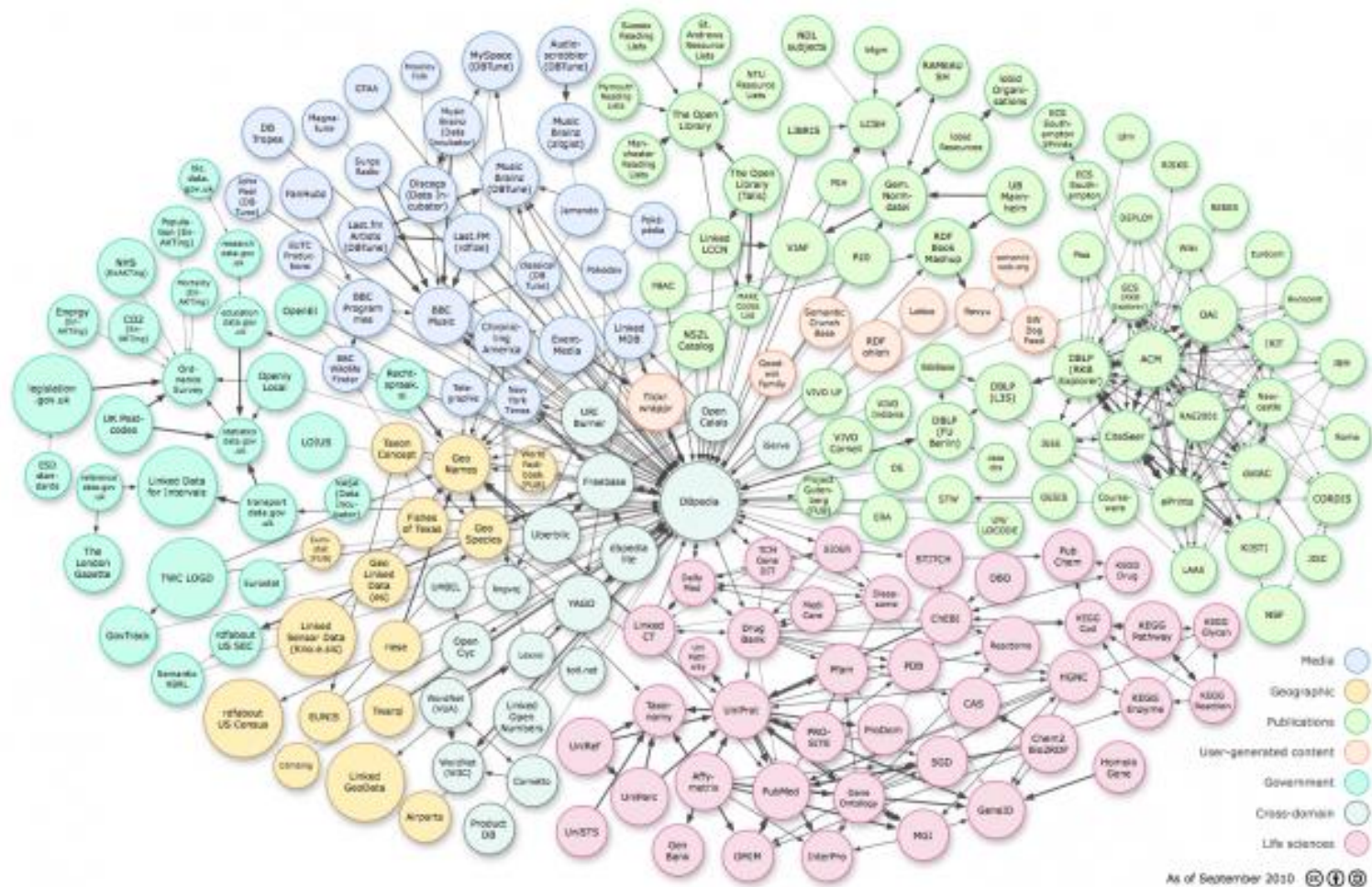


Exploit deeper knowledge

Important for missing labels problem

The Correlations are in fact structured

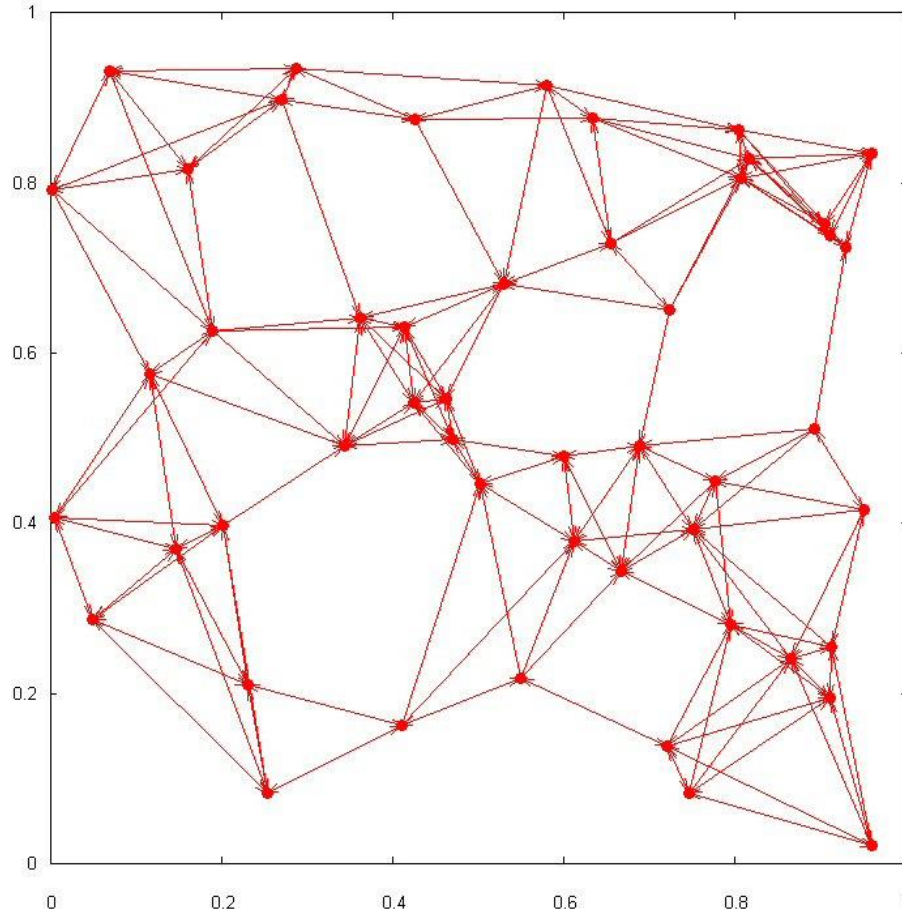
Label – Label correlations



Knowledge graph and Hierarchical tree graph can be used to model structured label – label correlations

Jia Deng *et al*,
ECCV 2014

Instance – Instance Correlations



NN Graph can be used to model structured instance instance correlations

Zhu et al. ACM MM 2010

Therefore, we want to formulate the problem so that we can

Exploit structured correlations, especially instance – instance correlations, in an efficient way.

Intuition

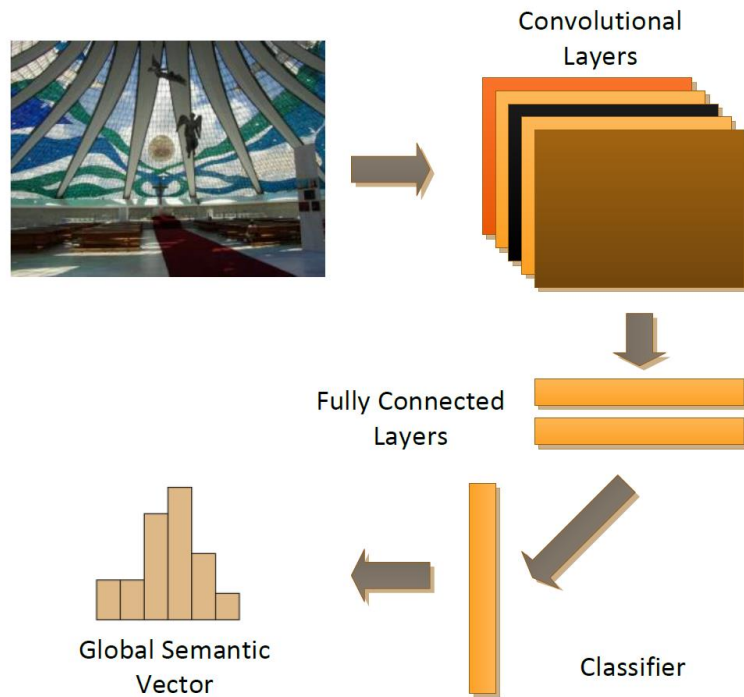
- The key to utilize structured instance-instance correlations is to make use of semantic correlations between images, as
- ***Semantically similar images should share similar labels***
 - *How to define a good semantic representation*
 - *How to construct a graph and incorporate it efficiently and effectively*

Semantic Feature Extraction

- Global Semantic Feature
- Local Semantic Feature

Semantic Feature Extraction

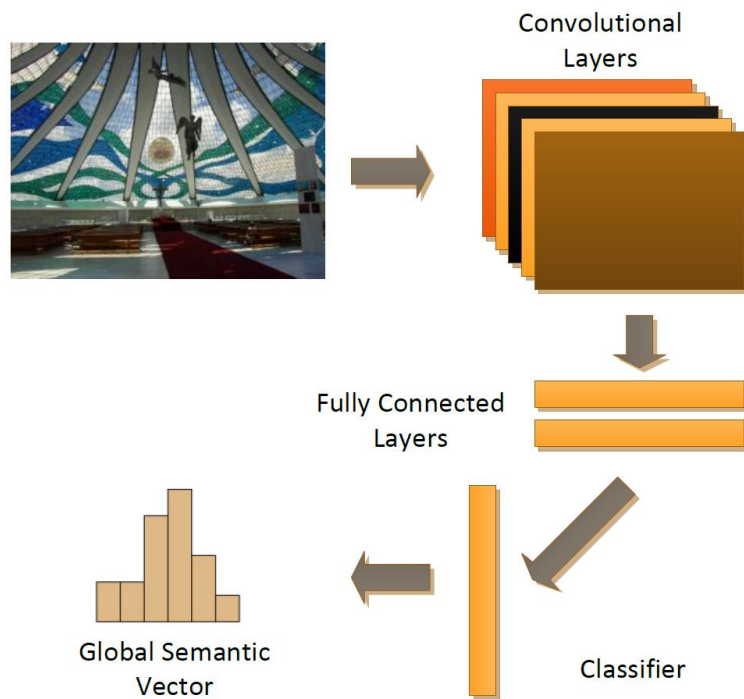
- **Global Semantic Descriptor**



Global semantic descriptors are extracted from relevant visual concepts from large scale datasets, e.g. ILSVRC, Places.

Semantic Feature Extraction

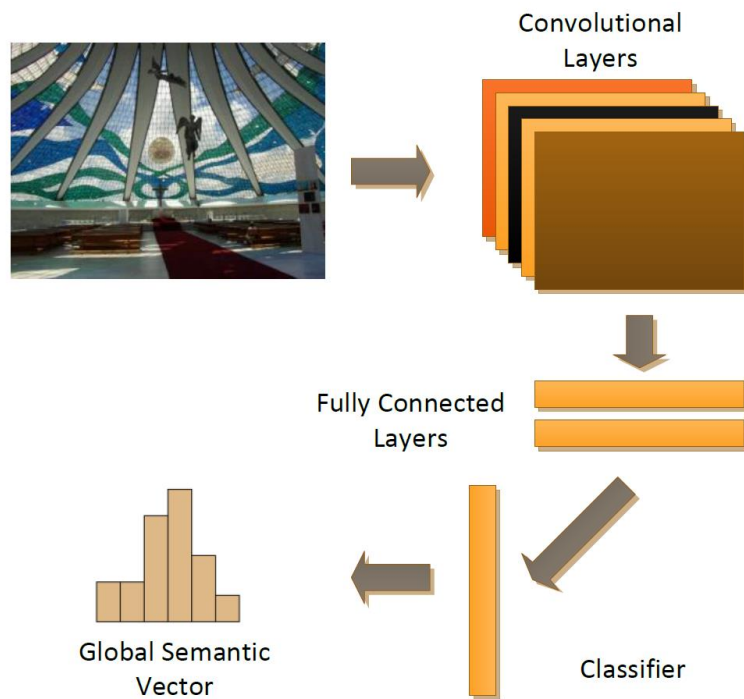
- **Global Semantic Descriptor**



Semantic Feature Extraction

- **Global Semantic Descriptor**

“what is the image in general”

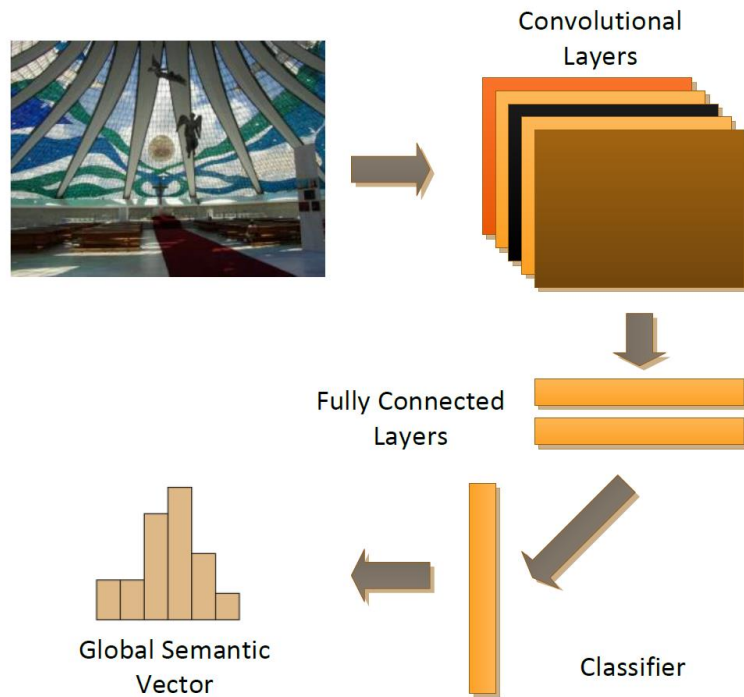


Semantic Feature Extraction

- **Global Semantic Descriptor**

“what is the image in general”

according to a large number of concepts developed in the general large-scale dataset



Local Semantic Descriptor

Local semantic descriptors are generated from labels of visual neighbours.



people, bottle



people, bottle



people



people, bottle



bed, lamp, night, painting, room
wall, window



bed, bedcover, curtain,
room, wall, window



bed, bedcover, curtain,
lamp, night, picture, side,
room, wall, window



bed, bedcover, curtain,
lamp, picture, room,
wall

Local Semantic Descriptor



people, bottle



people, bottle



people



people, bottle



bed, lamp, night, painting, room
wall, window



bed, bedcover, curtain,
room, wall, window



bed, bedcover, curtain,
lamp, night, picture, side,
room, wall, window



bed, bedcover, curtain,
lamp, picture, room,
wall

Local Semantic Descriptor

“what does the image specifically look like”.



people, bottle



people, bottle



people



people, bottle



bed, lamp, night, painting, room
wall, window



bed, bedcover, curtain,
room, wall, window



bed, bedcover, curtain,
lamp, night, picture, side,
room, wall, window



bed, bedcover, curtain,
lamp, picture, room,
wall

Graph Construction

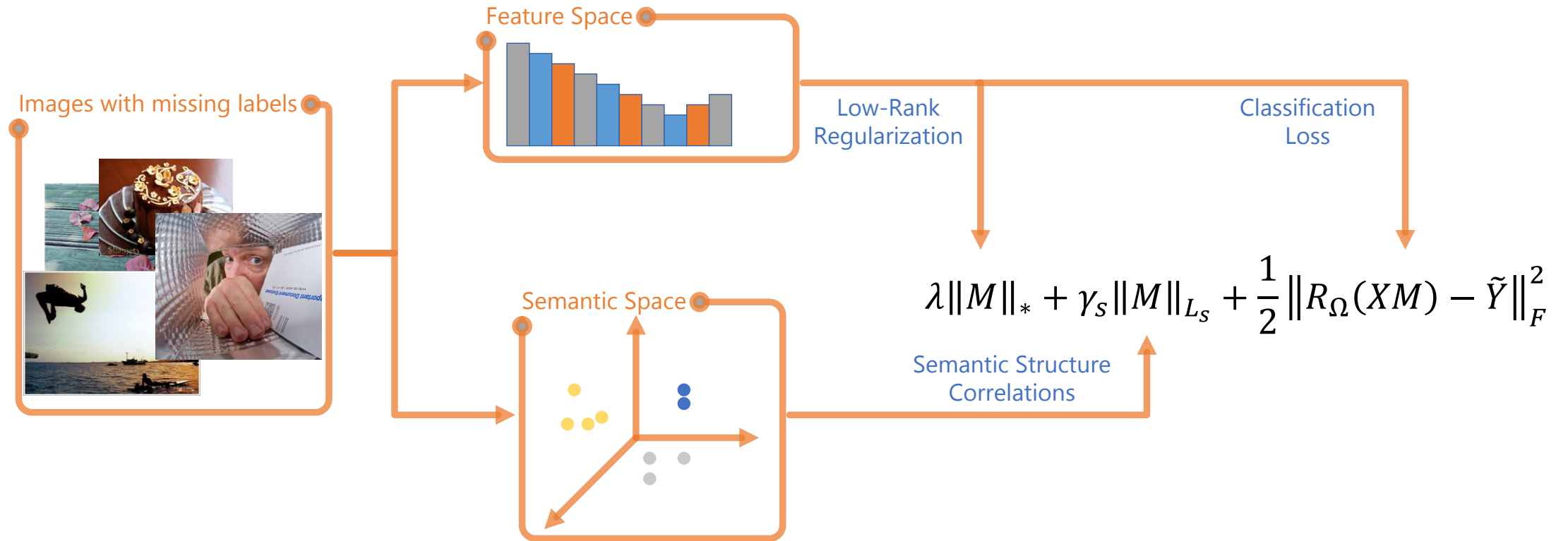
- Project images into semantic space
- Construct NN graph of semantic representations
- We can then incorporate structured instance - instance correlations with Laplacian regularization as shown in many related works.

$$\|M\|_{L_S} = \text{tr}(M^T X^T L_S X M)$$

Where L_S is the Laplacian of graph G_S in the semantic space.

- We can also easily add in label – label correlation if suitable graph is given

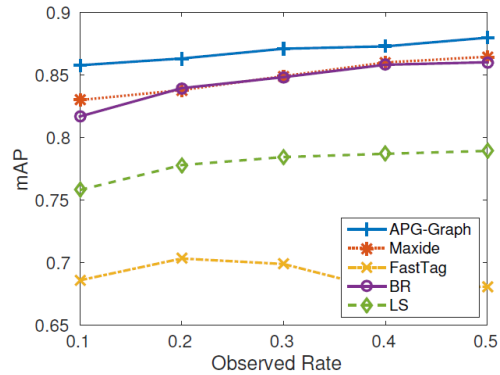
System Architecture



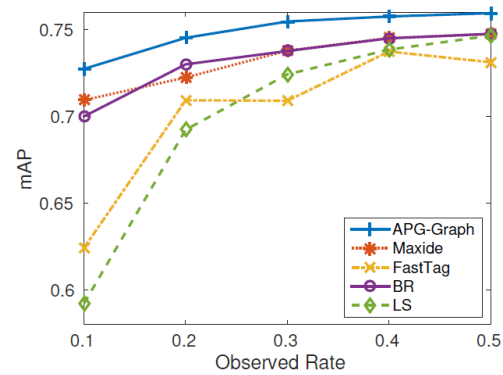
Experimental Results

- Datasets
 - FLIKR25K
 - PASCAL VOC2007
 - ESP GAME
 - IAPRTC-12

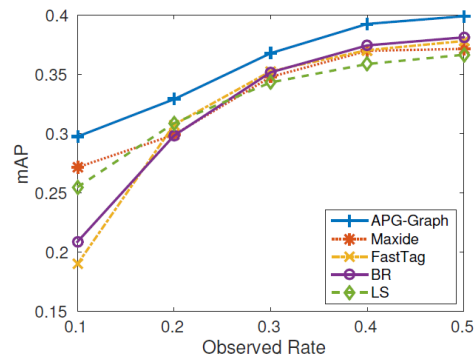
Experimental Results



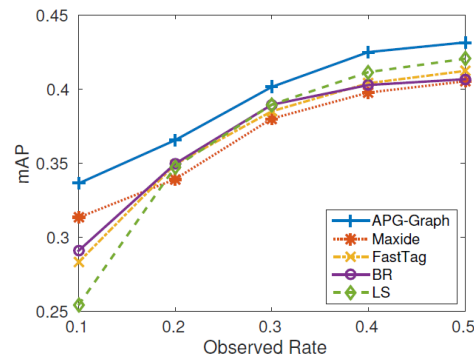
(a) VOC 2007



(b) FLICKR25K



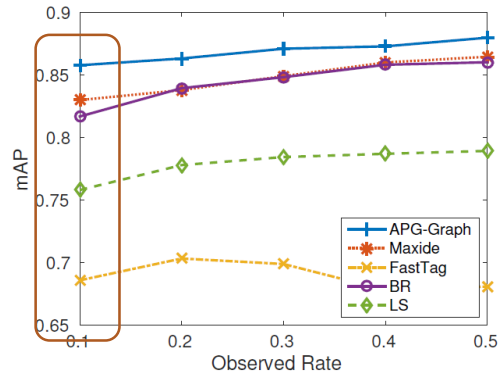
(c) ESP GAME



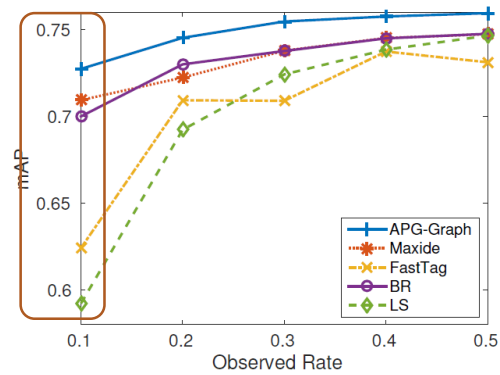
(d) IAPRTC-12

Comparison with several baseline methods on four multi-label datasets

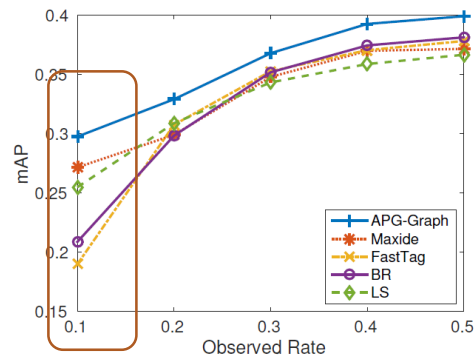
Experimental Results



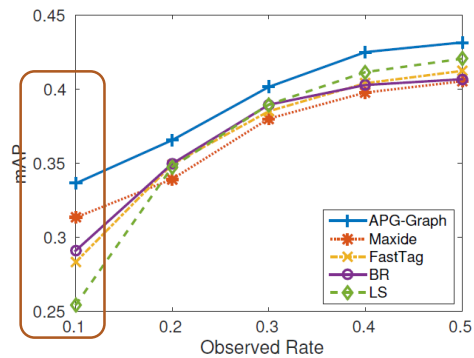
(a) VOC 2007



(b) FLICKR25K



(c) ESP GAME



(d) IAPRTC-12

Comparison with several baseline methods on four multi-label datasets

Experimental Results



car, person



person, bike



person, motor



boat, person



person, chair



man, hat, face,
black, yellow



white, metal, silver,
machine, water



white, car, tree
wheel, metal



tree, green, sky,
water, building



white, map, red,
chart, diagram

Some examples of labels
generated using our
method