

VSB²-Net: Visual-Semantic Bi-Branch Network for Zero-Shot Hashing

Summary

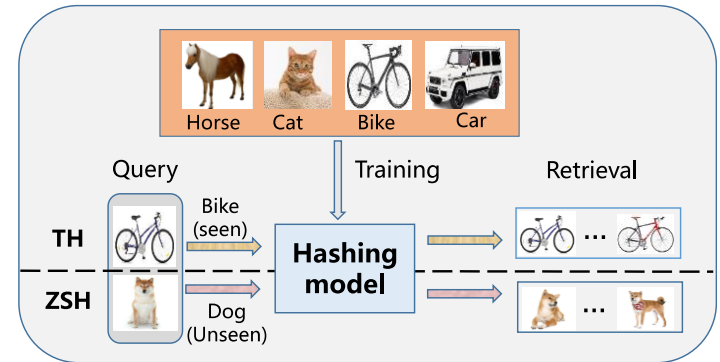
Problem: The existing methods mainly focused on optimizing the mapping between hash codes and semantic space, but ignored the core of the hash problem that generates discriminative hash codes.

Contributions:

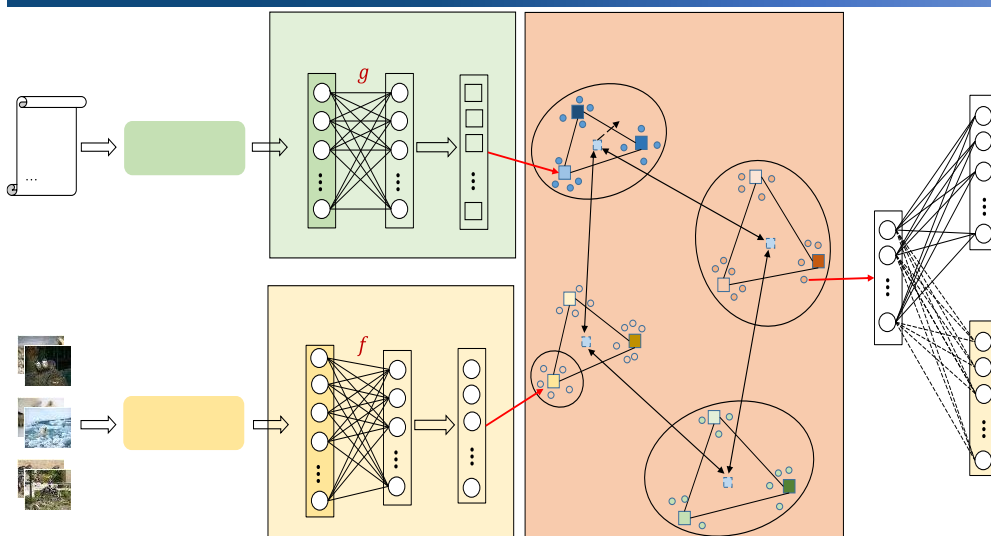
1. A novel inductive framework embeds the semantic space and visual feature space into the hamming space simultaneously.
2. The reconstructive semantic relationships are established to preserve local similarity relationships and explicitly enlarge the discrepancy.
3. A two-task architecture is directly employed to enhance the generalization and transfer abilities on unseen classes.

Problem Setting

Zero-shot hashing aims at learning hashing model from seen classes and the obtained model is capable of generalizing to unseen classes for image retrieval.



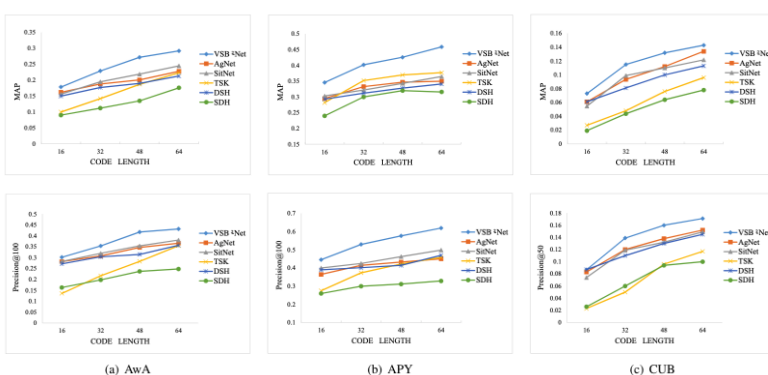
Framework



The architecture is a bi-branch network: the semantic similarity branch reconstructs the semantic space to hamming space, and the visual feature transfer branch embeds the visual features into the same hamming space simultaneously. In addition, task-driven regularization is composed of classification module and visual feature reconstruction module.

Main Result

Comparison of methods on public datasets.



Ablation Study

Effect of different number of class for training and testing on ImageNet dataset.



Selecting methods:

- (1) Wasserstein metric
- (2) Euclidean metric
- (3) Random selection

Method	AwA(MAP)			
	16 bits	32 bits	48 bits	64bits
Mini-SitNet	0.128	0.150	0.175	0.192
Mini-VSB ² -Net	0.151	0.178	0.220	0.241

Method	AwA(MAP)			
	16 bits	32 bits	48 bits	64bits
VSB ² -G	0.155	0.198	0.217	0.270
VSB ² -Net	0.175	0.230	0.272	0.293