Deep Learning of Binary Hash Codes for Fast Image Retrieval

Kevin Lin, Huei-Fang Yang, Jen-Hao Hsiao, Chu-Song Chen
Institute of Information and Science (IIS), Academia Sinica, Taipei, Taiwan
Research Center for Information Technology Innovation (CITI), Academia Sinica, Taipei, Taiwan
Yahoo! Taiwan

IEEE Conference on Computer Vision and Pattern Recognition,
DeepVision Workshop, 2015.
Extended version in arXiv Pre-print 1507.00101

Large-scale Image Search

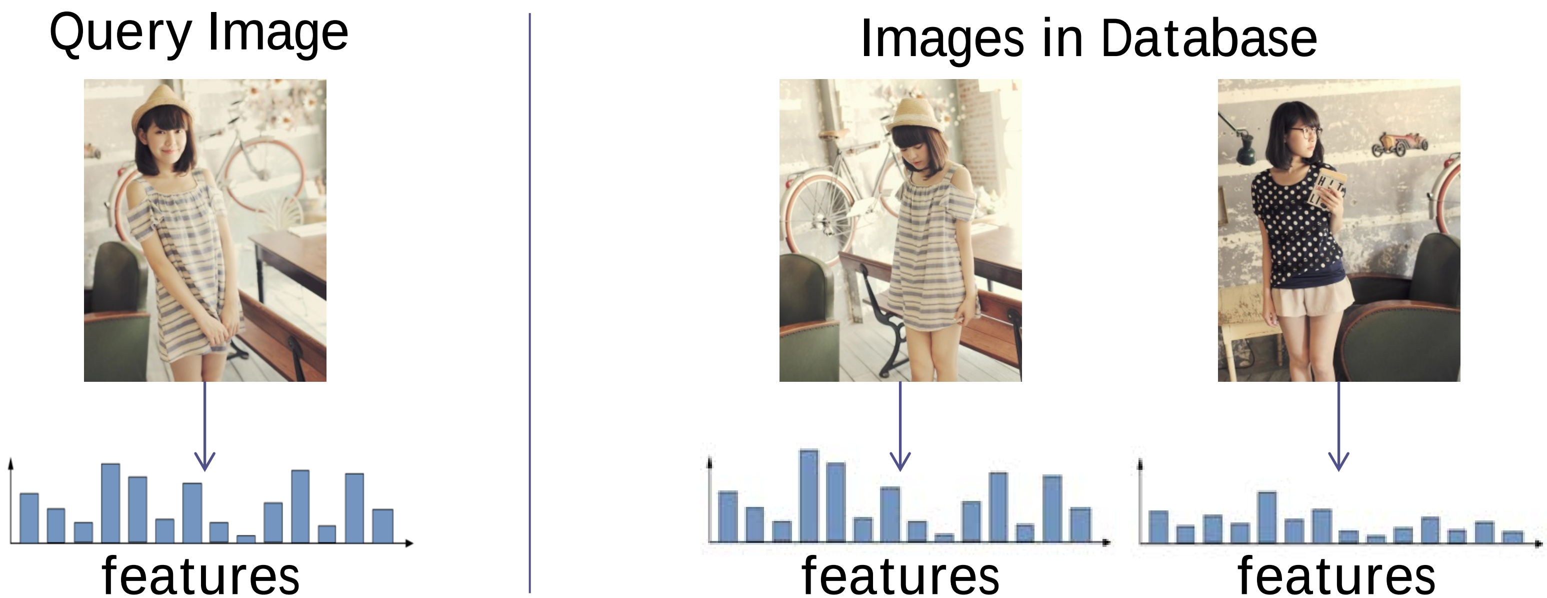
Query

Database

Top Retrieved Image

Search Strategy

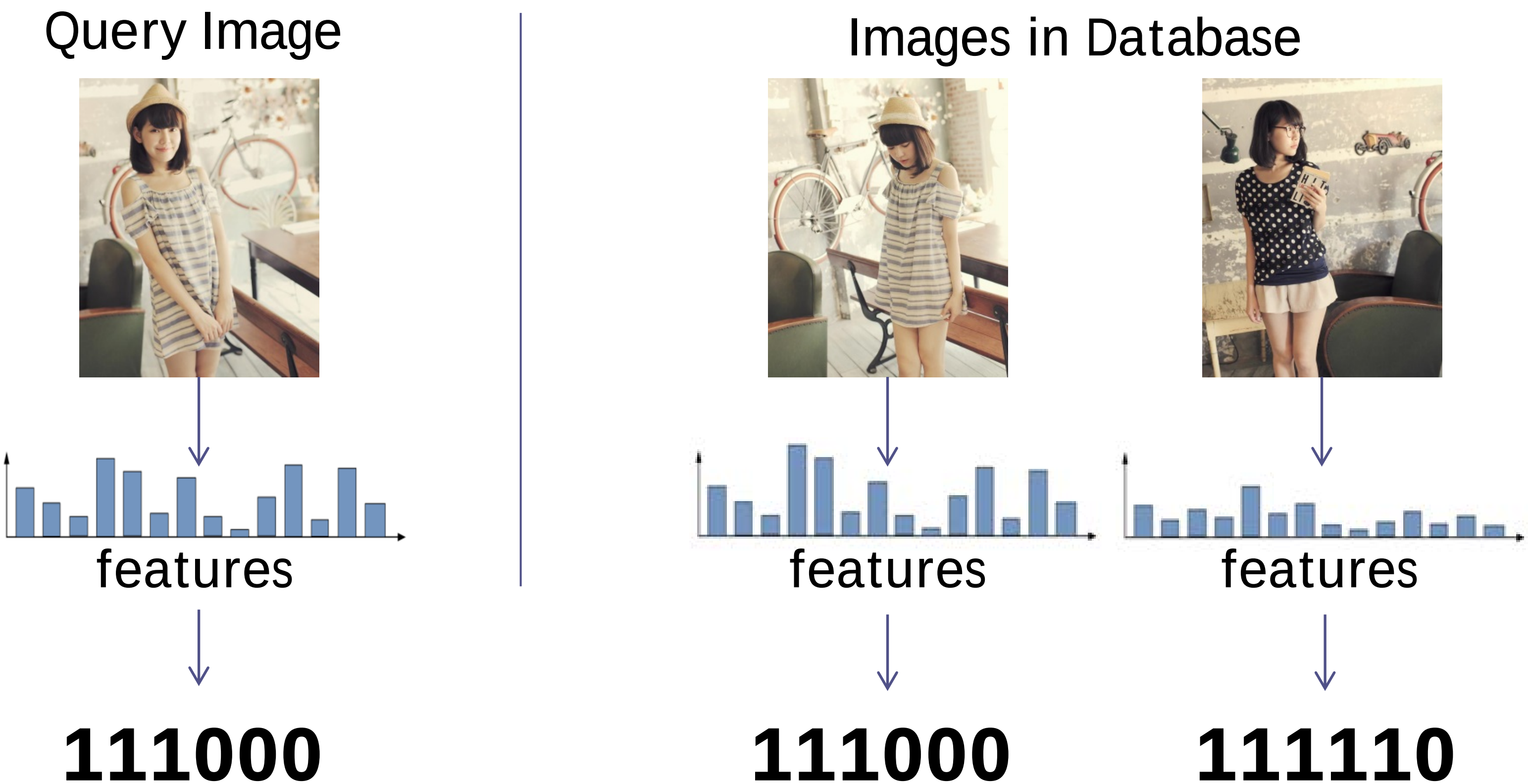
- Images are represented by features.



- Nearest neighbor search: neighbors of a point are determined by Euclidean distance.
- Challenge: How to efficiently search over millions or billions of images?

Solution: Binary Codes

- Images are represented by **binary codes**.



- Fast search can be carried out via Hamming distance measurement. (XOR operation)

Related works: Learning Binary Codes (1/2)

- Unsupervised learning
  - Use only training data, no label info.
  - Locality-Sensitive Hashing (LSH) [1]
  - Iterative quantization (ITQ) [2]
- Supervised learning
  - Use supervised info. (labels, pairwise similarities, …)
  - Binary reconstructive embedding (BRE) [3]
  - Minimal loss hashing (MLH) [4]

[1] A. Andoni and P. Indyk, "Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions," in Proc. FOCS, 2006.

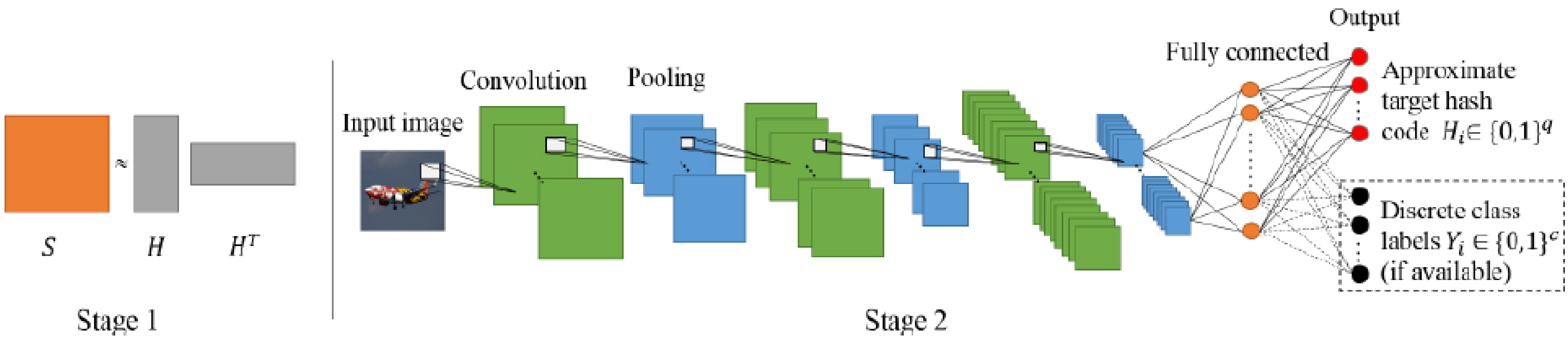
[2] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin, "Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 12, pp. 2916–2929, 2013.

[3] B. Kulis and T. Darrell, "Learning to hash with binary reconstructive embeddings," in Proc. NIPS, 2009.

[4] M. Norouzi and D. J. Fleet, "Minimal loss hashing for compact binary codes," in Proc. ICML, 2011.

Related works: Learning Binary Codes (2/2)

- Supervised deep Learning
  - Take advantage of deep neural network
  - Convolutional Neural Network Hashing (CNNH) [5]
  - Deep Neural Network Hashing (DNNH) [6]



[5] R. Xia, Y. Pan, H. Lai, C. Liu, and S. Yan. Supervised hashing for image retrieval via image representation learning. In Proc. AAAI Conference on Artificial Intelligence (AAAI), 2014.
[6] H. Lai, Y. Pan, Y. Liu, and S. Yan. Simultaneous feature learning and hash coding with deep neural networks. In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

Goal

- Can we take the advantage of deep CNN to achieve hashing?
- Can we generate the binary compact codes directly from the deep CNN?
- Solution: Supervised Semantic-preserving Deep Hashing (SSDH)

Approach

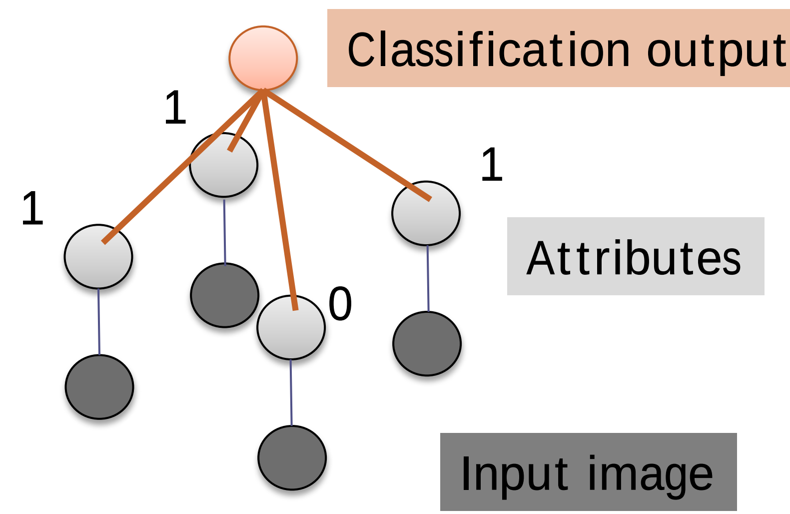
- Assume the classification outputs rely on a set of hidden attributes ***on*** and ***off***.

Bird

"furry"	0
"feather"	1
⋮	
"has head"	1
"has leg"	1

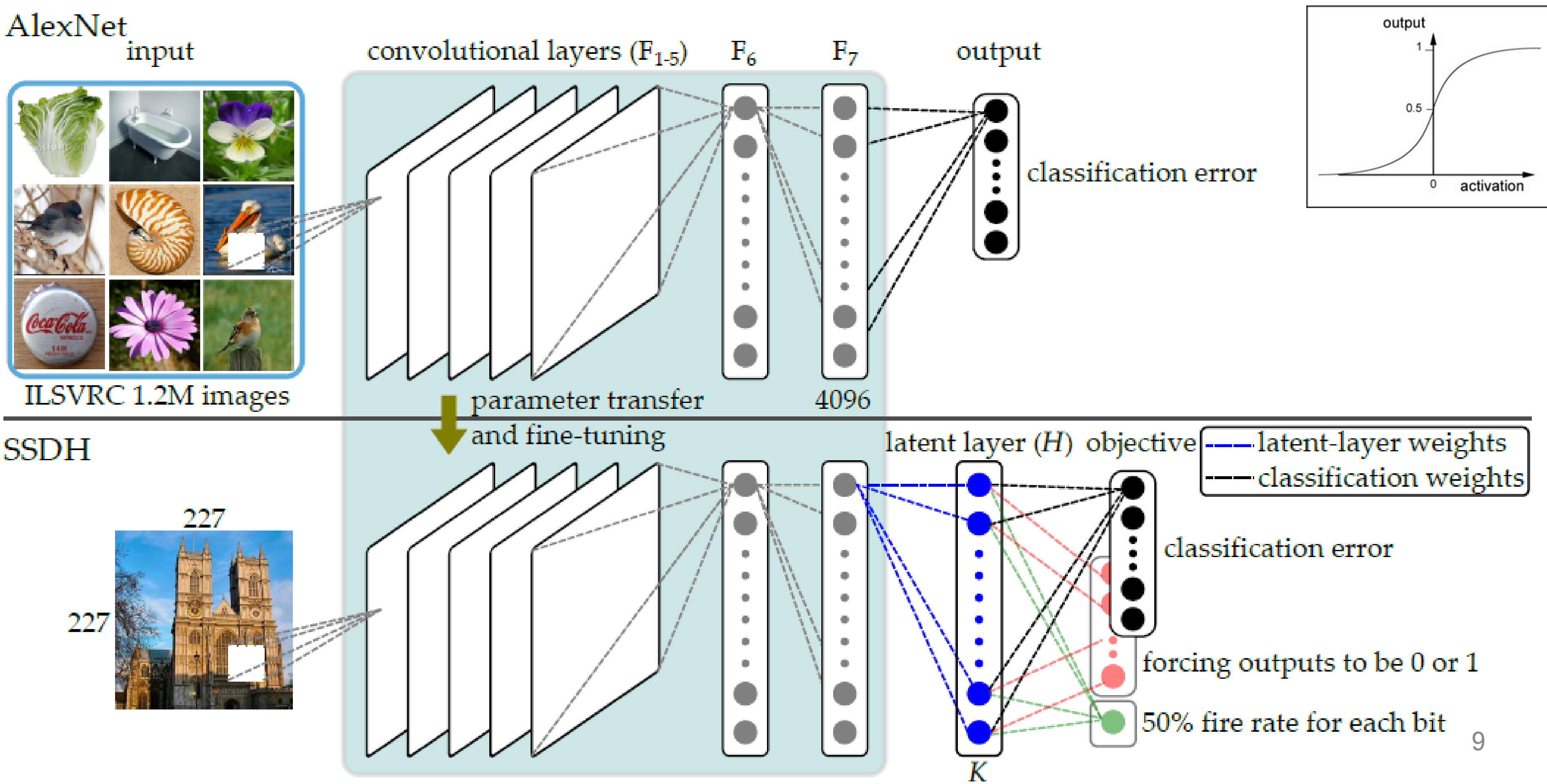
Cow

"furry"	1
"feather"	0
⋮	
"has head"	1
"has leg"	1

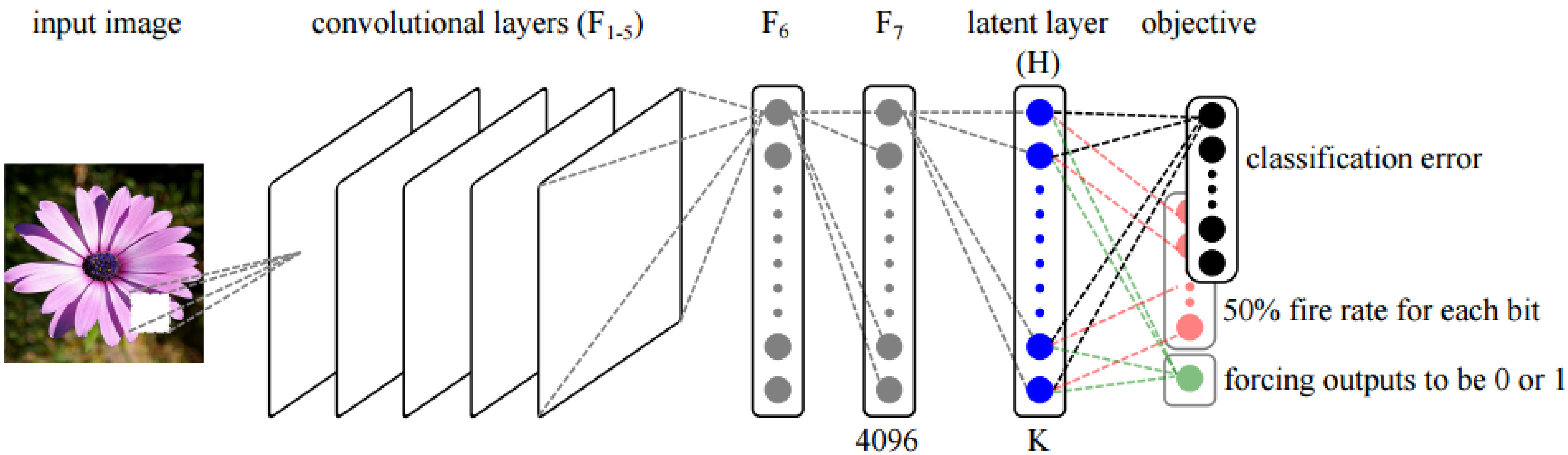


Approach

- We add the fully connected ***latent layer H*** between ***F7*** and ***F8***.
- The neurons in ***H*** are activated by sigmoid functions.



Approach

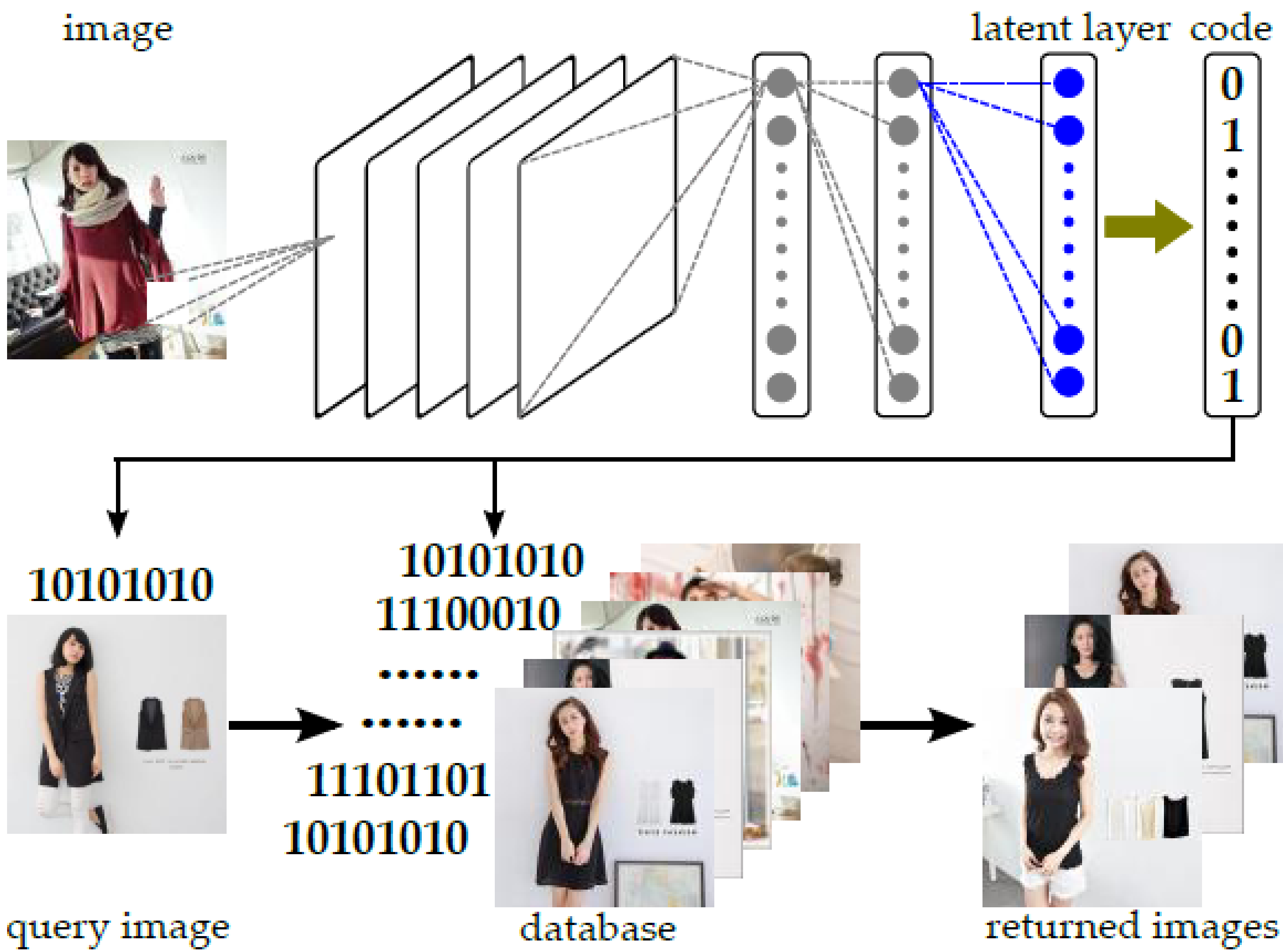


- Overall learning objective

$$\arg\min_{W} \quad \alpha \sum_{n=1}^{N} L(y_n, \hat{y}_n) + \lambda ||W||^2$$
$$- \beta \frac{1}{K} \sum_{n=1}^{N} ||a_n^H - 0.5\mathbf{e}||^2$$
$$+ \gamma \sum_{n=1}^{N} (\text{mean}(a_n^H) - 0.5)^2$$

Approach

- Compute binary codes for fast image retrieval

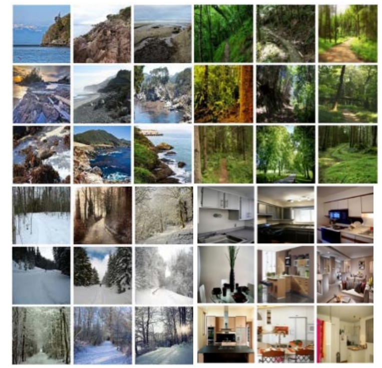


Experiments

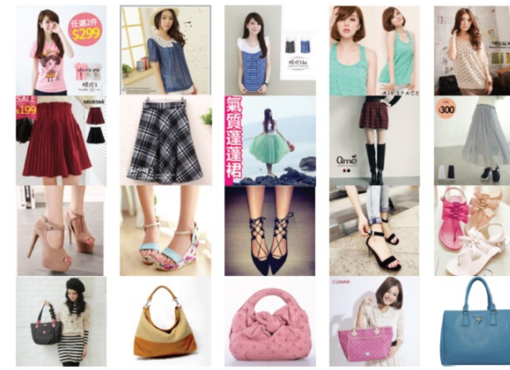
- Datasets

CIFAR10

MNIST



SUN397



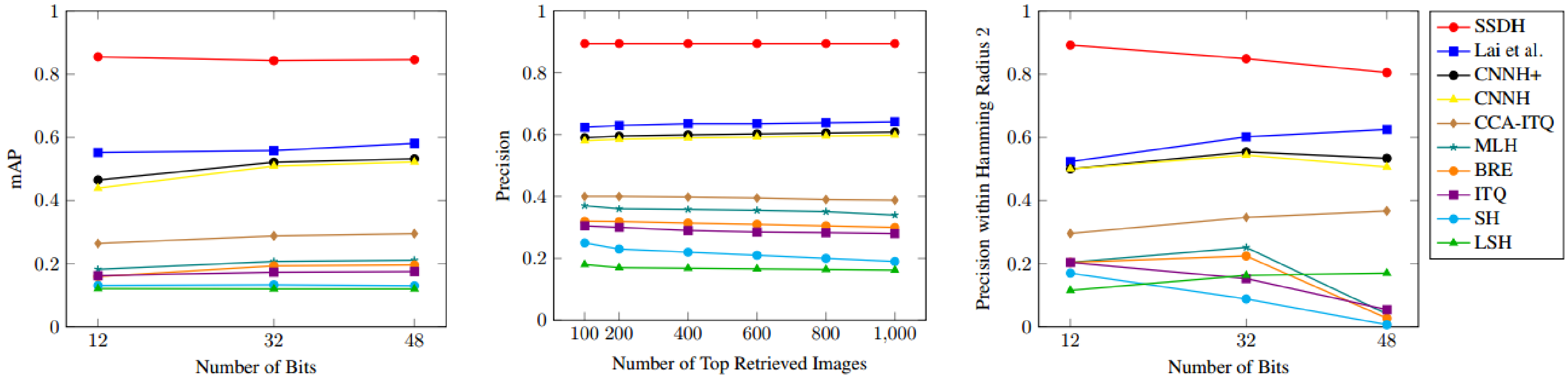
Yahoo-1M

TABLE
Statistics of datasets used in the experiments.

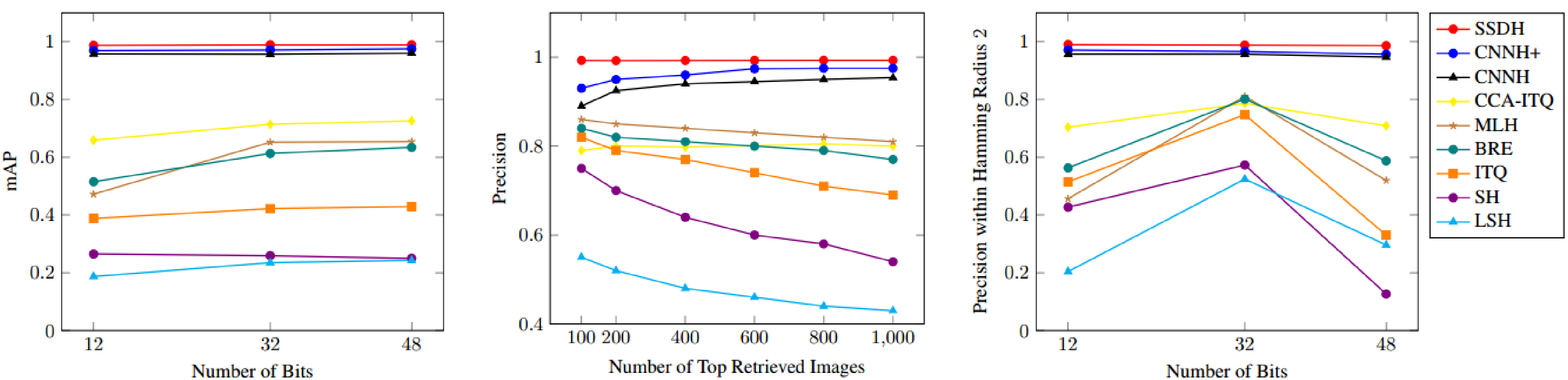
Dataset	# Labels	Training Set	Test Set
CIFAR-10	10	50,000	1,000
MNIST	10	60,000	10,000
SUN397	397	100,754	8,000
Yahoo-1M	116	1,011,723	112,363

Experiments

- CIFAR10



- MNIST



Experiments

Figure: Top 10 retrieved images from CIFAR-10 by varying the code length. Images with appearance more similar to the query are returned when the code length increases.

Experiments

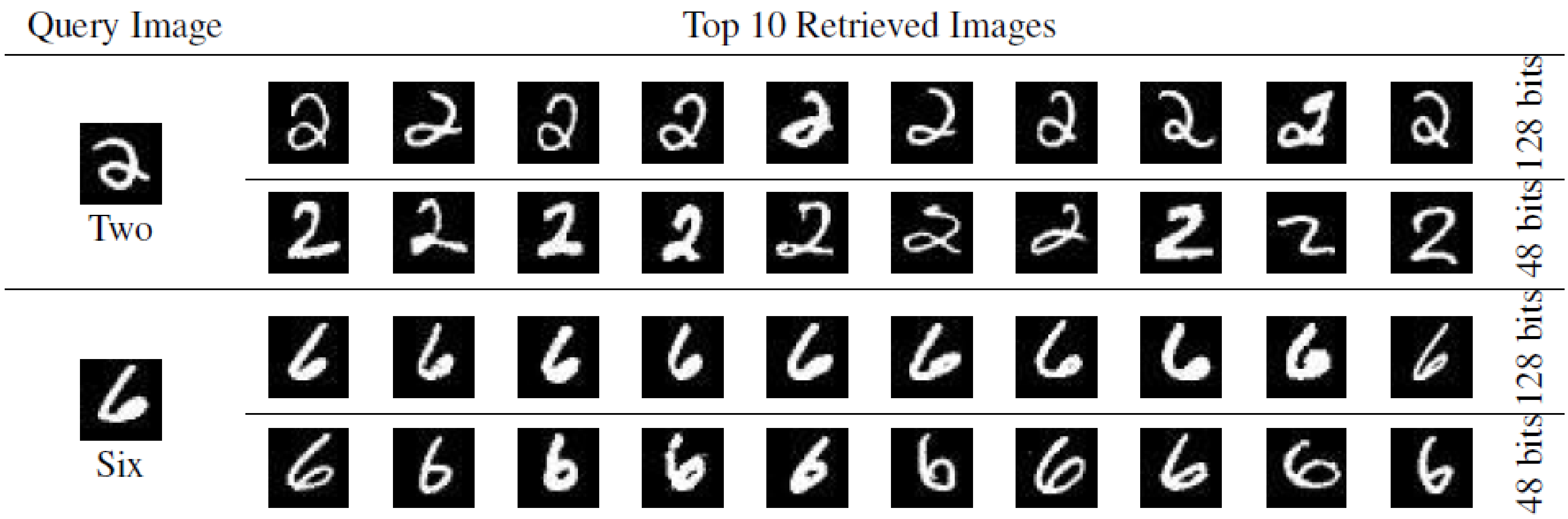


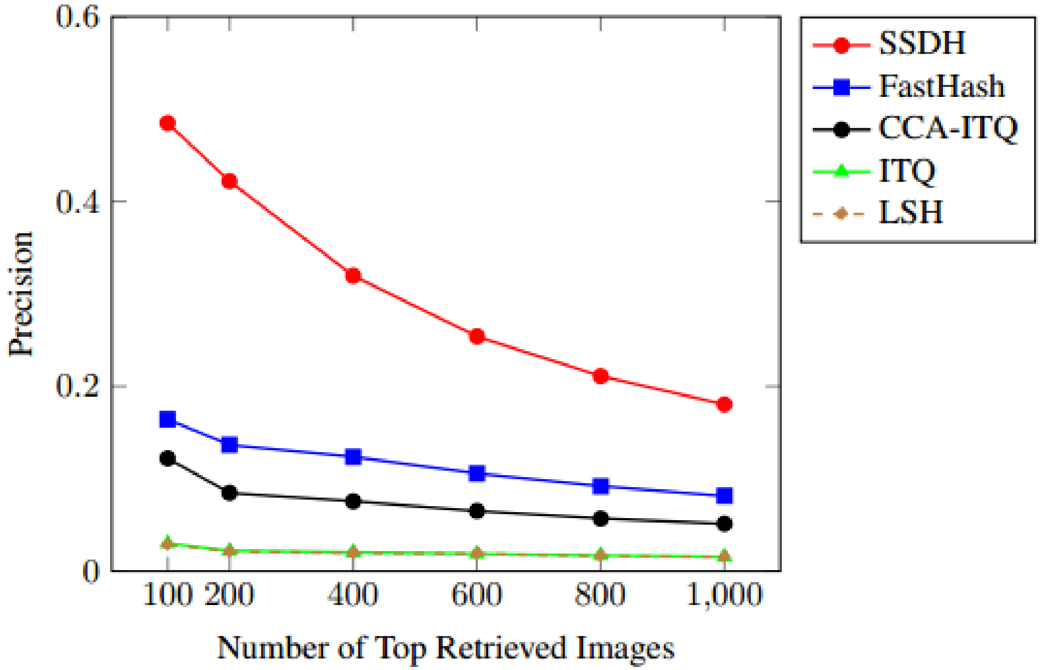
Figure: Top 10 retrieved images from MNIST by varying the code length. Images with appearance more similar to the query are returned when the code length increases.

Experiments

TABLE
Statistics of datasets used in the experiments.

Dataset	# Labels	Training Set	Test Set
CIFAR-10	10	50, 000	1, 000
MNIST	10	60, 000	10, 000
SUN397	397	100, 754	8, 000
Yahoo-1M	116	1, 011, 723	112, 363

- SUN397



[7]
[8]
[8]
[9]

Fig. 5. Precision curves with respect to different number of top retrieved samples on the SUN397 dataset when the 1024-bit hash codes are used in the evaluation.

[7] G. Lin, C. Shen, Q. Shi, A. van den Hengel, and D. Suter, “Fast supervised hashing with decision trees for high-dimensional data,” in Proc. CVPR, 2014
[8] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin, “Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 12, pp. 2916–2929, 2013.
[9] A. Andoni and P. Indyk, “Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions,” in Proc. FOCS, 2006

Experiments

TABLE
Statistics of datasets used in the experiments.

Dataset	# Labels	Training Set	Test Set
CIFAR-10	10	50,000	1,000
MNIST	10	60,000	10,000
SUN397	397	100,754	8,000
Yahoo-1M	116	1,011,723	112,363

- Yahoo-1M

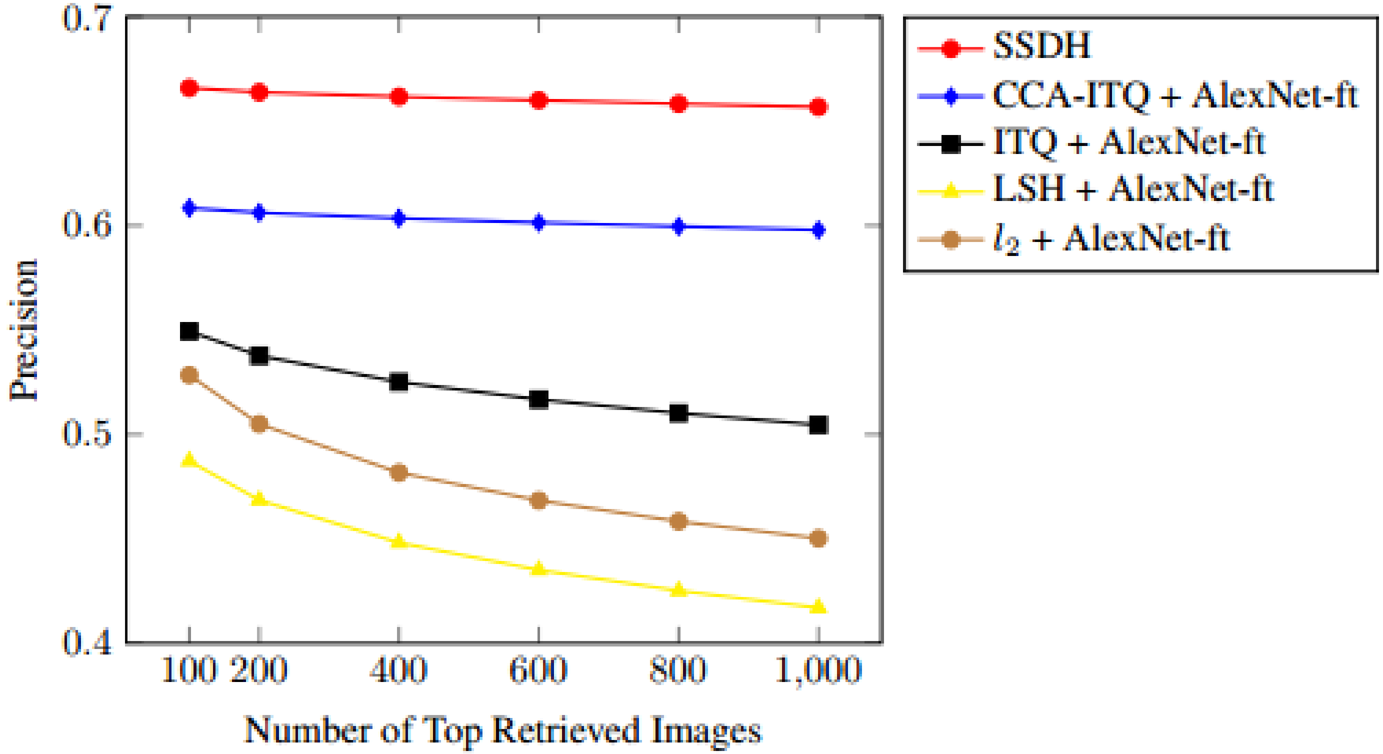


Fig. 6. Precision curves with respect to different number of top retrieved samples on the Yahoo-1M dataset when the 128-bit hash codes are used in the evaluation. AlexNet-ft denotes that the features from layer F_7 of AlexNet fine-tuned on Yahoo-1M are used in learning hash codes.

Experiments

Query	Top 5 Retrieved Image

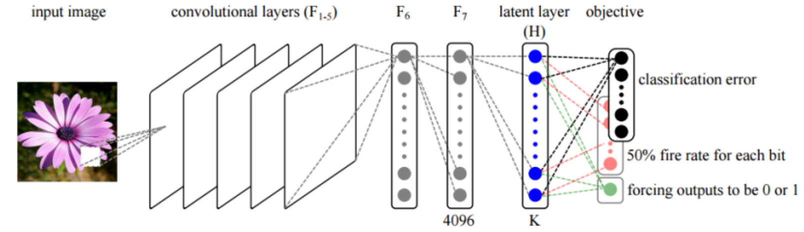
Ours

AlexNet

Ours

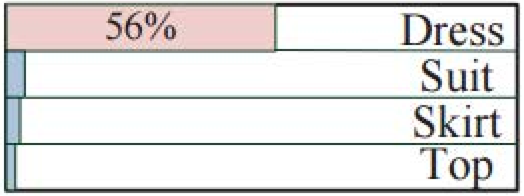
AlexNet

Experiments

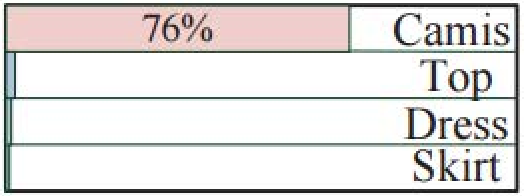


- Image classification results

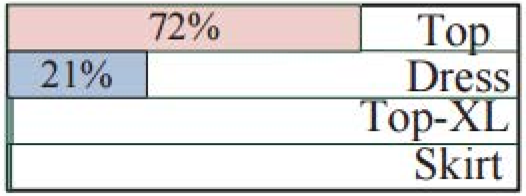
Dress



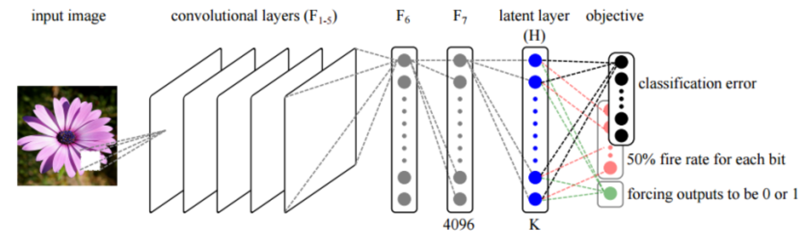
Camis



Top



Experiments



- Image classification results

Experiments

- Image classification results

TABLE 3
Classification accuracy of various methods on CIFAR-10.

Method	Accuracy (%)
Stochastic Pooling [64]	84.87
CNN + Spearmint [65]	85.02
AlexNet + Fine-tuning	88.31
NIN + Dropout [66]	89.59
NIN + Dropout + Augmentation [66]	91.19
SSDH w/ 12-bit codes	88.75
SSDH w/ 32-bit codes	89.50
SSDH w/ 48-bit codes	89.59

TABLE 4
Classification accuracy of various methods on MNIST.

Method	Accuracy (%)
Stochastic Pooling [64]	99.53
AlexNet + Fine-tuning	99.16
NIN + Dropout [66]	99.53
SSDH w/ 12-bit codes	99.25
SSDH w/ 32-bit codes	99.09
SSDH w/ 48-bit codes	99.16

TABLE 5
Classification accuracy of various methods on SUN397.

Method	Accuracy (%)
AlexNet + Fine-tuning	47.94
Cascade fine-tuned CNN [67]	46.87
MOC-CNN [68]	51.98
SSDH w/ 48-bit codes	46.16
SSDH w/ 128-bit codes	49.13
SSDH w/ 1024-bit codes	49.03

TABLE 6
Classification accuracy of various methods on Yahoo-1M.

Method	Accuracy (%)
AlexNet + Fine-tuning	71.28
SSDH w/ 128-bit codes	72.38

[64] M. D. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks," in Proc. ICLR, 2013.
[65] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," in Proc. NIPS, 2012.
[66] M. Lin, Q. Chen, and S. Yan, "Network in network," in Proc. ICLR, 2014.
[67] Z. Jie and S. Yan, "Robust scene classification with cross-level LLC coding on CNN features," in Proc. ACCV, 2014
[68] Y. Gong, L. Wang, R. Guo, and S. Lazebnik, "Multi-scale orderless pooling of deep convolutional activation features," in Proc. ECCV, 2014.

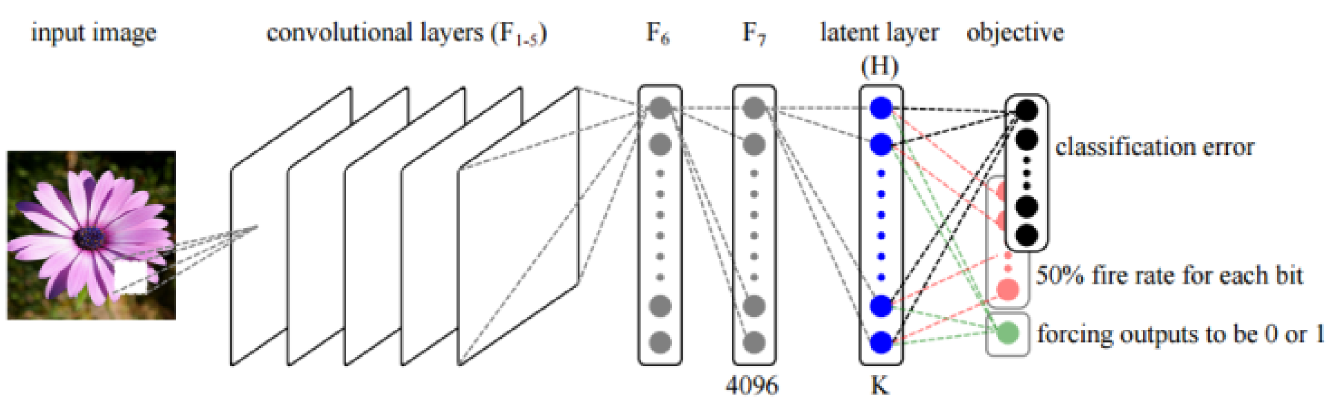
Experiments

Descriptor	Measurement	Time
CNN-fc7-4096	Euclidean distance	22.6 µs
SSDH-64	Hamming distance	23.0 ps

- Search with 64-bit codes is approximately ~982,600x faster than traditional exhaustive search with 4096-dimensional deep features.

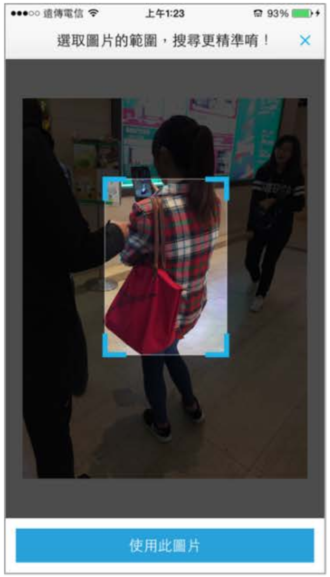
Conclusion

- SSDH constructs hash functions as a latent layer between the feature layer and classification layer in a network.
- SSDH jointly learn binary codes, features, and classification by optimizing the parameters of the network with the proposed objective function.
- SSDH is scalable to large scale search.



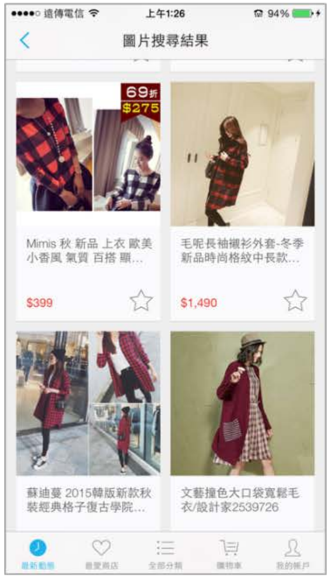
Application: Mobile clothing search

- The technology has been integrated in YAHOO!奇摩超級商城



Take photo

Select category



Similar clothing

Buy it!

- Download YAHOO!奇摩超級商城 Apps

View in iTunes

Install

Thank you!

Download our codes and models at

https://github.com/kevinlin311tw/caffe-cvprw15

GitHub