“Deep Supervised Hashing for Fast Image Retrieval”

Haomiao Liu
Ruiping Wang
Shiguang Shan
Xilin Chen

PRESENTED BY –
Aarathi M (20173018)
ROHIT TIWARI (20162085)

Abstract

- Use of hashing to learn *compact binary codes for highly efficient image retrieval on large-scale datasets*
- Uses the power of CNN
- Supervised Learning
    - ➢ I/P : pairs of images (similar/dissimilar), discrete values (eg. +1/-1)
    - ➢ O/P: K-bit binary codes and images similar to buried images

Walk-through

- Abstract
- Introduction
- Approach
- Loss Function
- Relaxation
- Implementation
- Conclusion

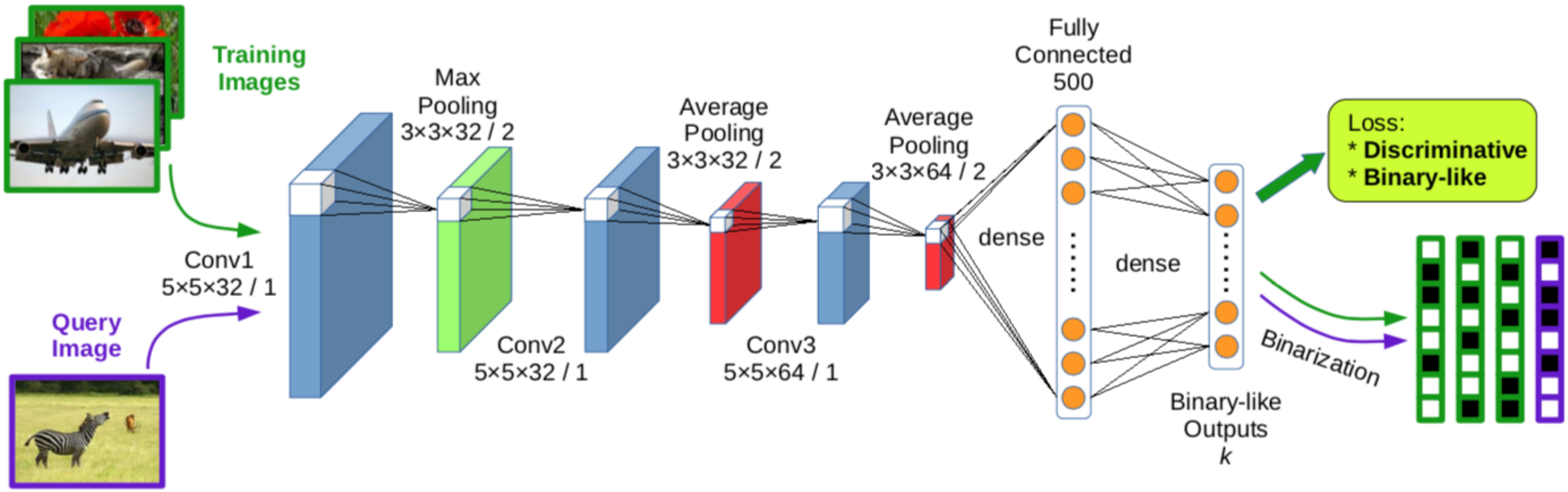
Introduction

- Goal : learning compact binary codes for images such that
  - similar images should be encoded to similar binary codes in Hamming space
  - the binary codes should be computed efficiently
- Use of CNNs to capture both visual as well as semantic similarity of images.

Approach

- Use CNNs to break out the limitations of both hand- crafted features and linear models.
- Train the CNN using image pairs and the corresponding similarity labels.
- Quantize the CNN outputs to generate binary codes for new-coming images.

Approach


Training Images
Query Image
Conv1 5×5×32 / 1
Max Pooling 3×3×32 / 2
Conv2 5×5×32 / 1
Average Pooling 3×3×32 / 2
Conv3 5×5×64 / 1
Average Pooling 3×3×64 / 2
dense
Fully Connected 500
dense
Binary-like Outputs k
Loss:
* Discriminative
* Binary-like
Binarization

Loss Function

- Loss Function is defined as-

$$L(\mathbf{b}_1, \mathbf{b}_2, y) = \frac{1}{2}(1-y)D_h(\mathbf{b}_1, \mathbf{b}_2) + \frac{1}{2}y \max(m - D_h(\mathbf{b}_1, \mathbf{b}_2),\ 0)$$
$$s.t.\ \ \mathbf{b}_j \in \{+1, -1\}^k,\ j \in \{1, 2\}$$

where,

- $\mathbf{b}_1$, $\mathbf{b}_2$ belong to $\{+1, -1\}^k$
- y = 0 if they are similar, and y = 1 otherwise
- $D_h(.,.)$ denote the Hamming distance between two vectors
- m > 0 is a margin threshold parameter

Relaxation

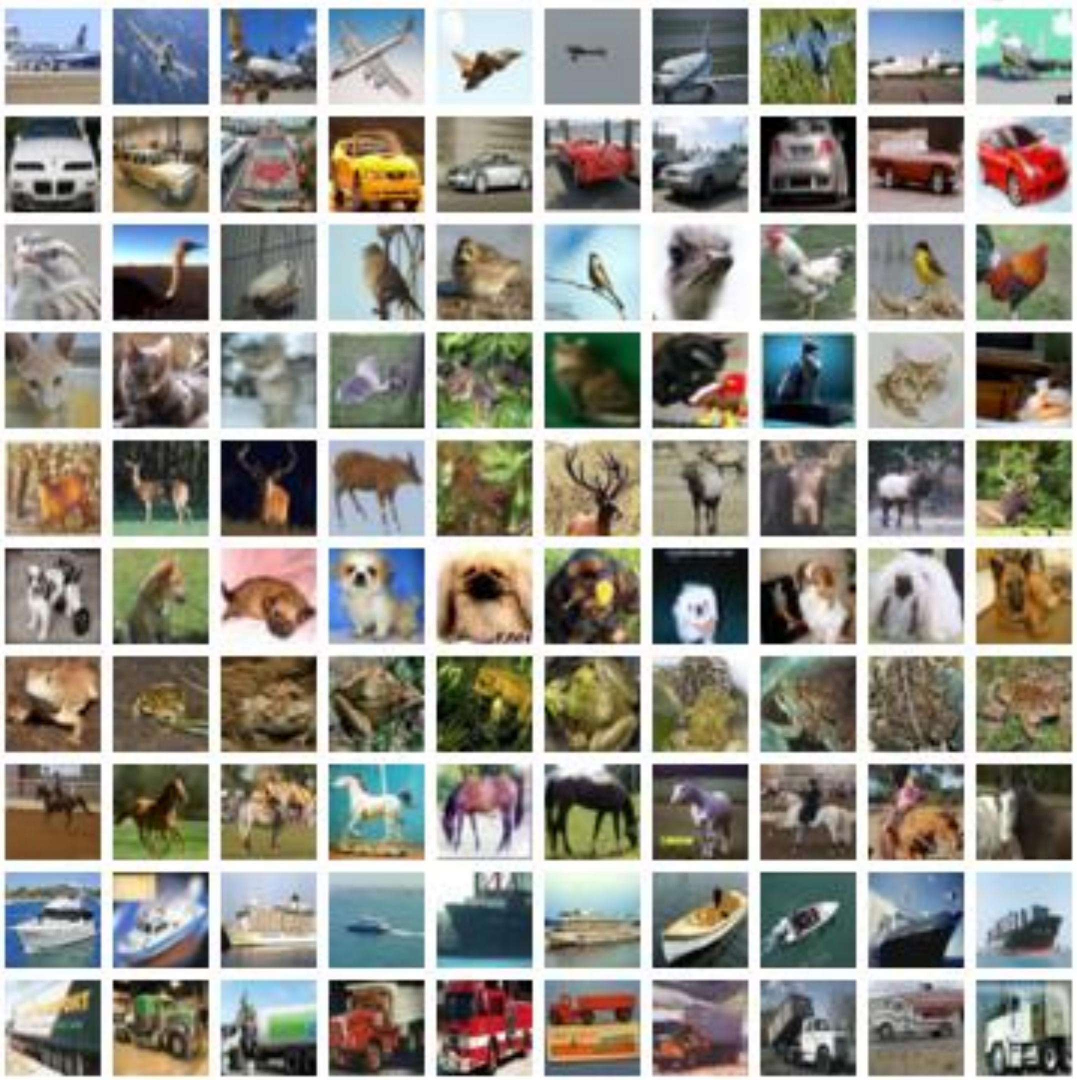
- Hamming distance replaced by Euclidean distance
- Additional regularizer added to replace the binary constraints
- Relaxed Loss Function -

$$
\begin{aligned}
L_r(\mathbf{b}_1, \mathbf{b}_2, y) =& \frac{1}{2}(1-y)||\mathbf{b}_1 - \mathbf{b}_2||_2^2 \\
&+ \frac{1}{2}y \ \max(m - ||\mathbf{b}_1 - \mathbf{b}_2||_2^2, \ 0) \\
&+ \alpha(|| \, |\mathbf{b}_1| - \mathbf{1}||_1 + || \, |\mathbf{b}_2| - \mathbf{1}||_1)
\end{aligned}
$$

where,

- **1** is the
- $||.||_1$ is
- |.| is the element wise absolute value operation
- alpha is the weighting parameter

Dataset

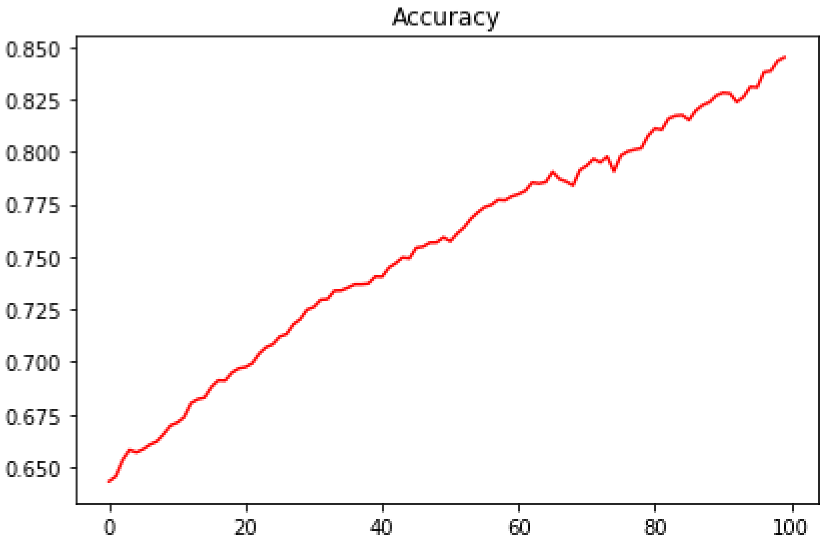


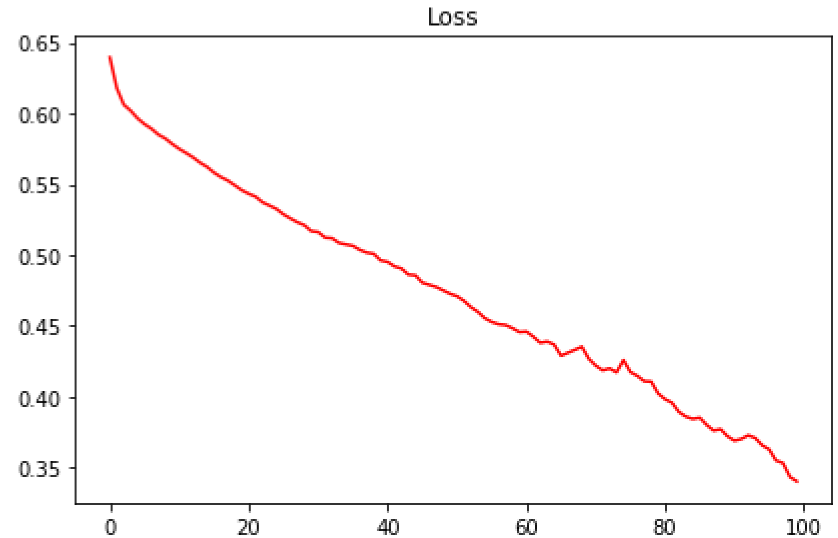
- CIFAR-10 dataset
- 50,000 images – 10 classes
- size 3*32*32.

- Image pairs generated randomly
- Images of same class are similar images
- Images of different classes are dissimilar images

Results

- 12 bit binary outputs
- The outputs of Model-1 for two images (b1-b2) are merged (abs(b1-b2))
- This model is trained with the training dataset created earlier.
- No regularisation
- Number of Image pairs: 200000
- Number of epochs: 100
- Batch Size: 200





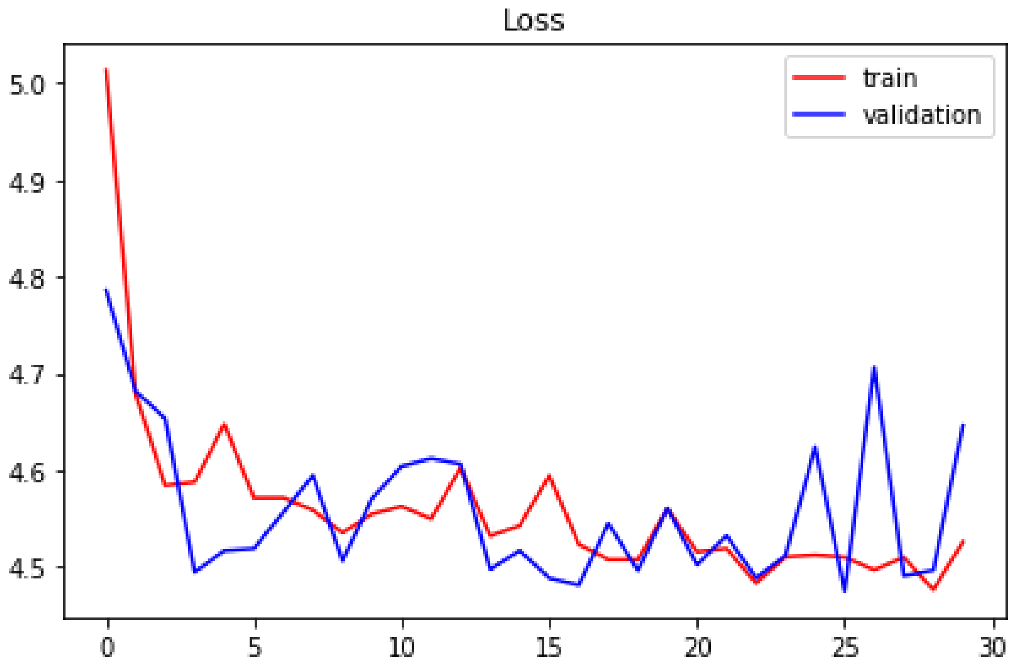
Results

- 100000000001 Frog
- 010011110000 Truck
- 011011110010 Truck
- 100000000001 deer
- 010001110000 automobile
- 111011000000 automobile
- 100100001101 bird
- 100100000100 horse
- 111010010010 ship
- 100000000000 cat
- 100000000001 deer

Precision = 0.7176
Recall = 0.7268
Average precision = 0.6942

Results

- 12 bit binary outputs
- Euclidian distance
- Margin (m) for loss calculation = 24
- L1 Regularization parameter = 0.01
- Number of Image pairs: 200000
- Number of epochs: 30
- Batch Size: 200



References

- H. Liu, R. Wang, S. Shan and X. Chen, "Deep Supervised Hashing for Fast Image Retrieval," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 2064-2072.*
- T.-S. Chua, J. Tang, R. Hong, H. Li, Z. Luo, Y. Zheng, "Nus-wide: A realworld web image database from national university of singapore", Proceedings of the ACM International Conference on Image and Video Retrieval, pp. 48, 2009.
- J. Deng, N. Ding, Y. Jia, A. Frome, K. Murphy, S. Bengio, Y. Li, H. Neven, H. Adam, "Large-scale object classification using label relation graphs", ECCV 2014, pp. 48-64, 2014.

ThankYou