Deep Residual Hashing Network for Image Retrieval

Edwin Jimenez-Lepe[1] and Andres Mendez-Vazquez[1]

CINVESTAV Guadalajara
[eejimenez, amendez]@gdl.cinvestav.mx
http://www.gdl.cinvestav.mx

Abstract. Conventional methods in Content-Based Image Retrieval use hand-crafted visual features as input but sometimes such feature vectors do not preserve the similarity between images. Taking advantage of the improvements in the Convolutional Neural Networks (CNN) area we propose a Deep Residual Hashing Network (DRHN) based in the work of [1] that generates binary hash codes based on the features learned.
The base of our model is a residual block which is formed by the following operations: Convolution, Rectified Linear Unit (ReLU), Batch normalization (BN) [2] and Element-wise addition. A Residual Group is the join of n Residual Blocks. The architecture of the proposed model is the next: Input $\rightarrow$ (Convolution $\rightarrow$ ReLU $\rightarrow$ BN) $\rightarrow$ six residual blocks $\rightarrow$ Average Pooling Layer (APL) $\rightarrow$ Hash Layer $\rightarrow$ Fully Connected Layer with Softmax, where the APL calculates the average of all values in a channel and Hash Layer length is denoted by h.
We exploit a previous proposed idea in the field [3]: in a supervised manner the binary codes can be learned adding an extra hidden layer to represent the main features that identifies the classes in a database, hence to generate the binary code of h bits related to an image we binarize the activation of the Hash Layer according to some threshold t.
The experimental results outperforms the state-of-the-art hashing algorithms on the CIFAR-10 dataset, using a DRHN with $n = 15$ and $h = 48$ we obtained a mean average precision (mAP) of 92.91 in the image retrieval task, previous best known result for the same number of bits was 89.73 [3]

Keywords: Convolutional Neural Networks, Content-Based Image Retrieval, Computer Vision, Deep Learning

1. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778 (2016)
2. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
3. Lin, K., Yang, H.F., Hsiao, J.H., Chen, C.S.: Deep learning of binary hash codes for fast image retrieval. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. pp. 27–35 (2015)