Restricted Boltzmann Machine for Binary Patterns Aggregation for Image Object Recognition

Rafal Kapela, Szymon Sobczak, Dariusz Pazderski, Krzysztof Kozlowski, Aleksandra Swietlicka
Poznan University of Technology
rafal.kapela@put.poznan.pl, szymon.k.sobczak@doctorate.put.poznan.pl, dariusz.pazderski@put.poznan.pl, krzysztof.kozlowski@put.poznan.pl, aleksandra.swietlicka@put.poznan.pl

Kevin McGuinness, Noel O'Connor
Dublin City University
kevin.mcguinness@dcu.ie, noel.oconnor@dcu.ie

Abstract

The article presents a new approach for image object recognition and object classification taking advantage of binary local image descriptors. Such descriptors are known to be a relatively fast and effective solution to finding and describing image patches around previously found key-points. The most popular examples are Fast Retina Keypoint (FREAK), Binary Robust Invariant Scalable Keypoints (BRISK), Learned Arrangements of Three Patch Codes (LATCH). All of them exploit binary response as an image patch descriptor which can ensure fast operation. Simultaneously, this technique suffers from the lack of efficient aggregation methods.

Restricted Boltzmann Machine is a bi-directional neural network that can learn probability distribution of a binary input pattern. In our approach, RBMs are applied to transform a binary string to a vector of continuous values which can encode a set of binary patterns, each corresponding to the given previously calculated binary descriptor. For extraction of the object properties a simple sliding window approach can be employed. In order to find objects candidates in an image, we also introduce a Selective Search algorithm which combines the strength of both an exhaustive search and segmentation. For research purposes, we take advantage of our own application of Restricted Boltzmann Machine written in C++. The novelty considered in this paper is related to presentation of recognition results of our algorithm as well as proposition of the new implementation of the RBM using CUDA for GPU accelerated computing.

Based on obtained experimental results we can state that our technique using local image descriptors provides significantly better performance in comparison to simpler approaches referring to nearest neighbor and bag of visual words.

References

1. A. Fischer and C. Igel. An Introduction to Restricted Boltzmann Machines. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Vol. 7441 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012, pp. 14-36..