

Spatial Pooling of Heterogeneous Features for Image Classification

Abstract:

In **image** classification tasks, one of the most successful algorithms is the bag-of-features (BoFs) model. Although the BoF model has many advantages, such as simplicity, generality, and scalability, it still suffers from several drawbacks, including the limited semantic description of local descriptors, lack of robust structures upon single visual words, and missing of efficient spatial weighting. To overcome these shortcomings, various techniques have been proposed, such as extracting multiple descriptors, spatial context modeling, and interest region detection. Though they have been proven to improve the BoF model to some extent, there still lacks a coherent scheme to integrate each individual module together. To address the problems above, we propose a novel framework with spatial pooling of complementary features. Our model expands the traditional BoF model on three aspects. First, we propose a new scheme for combining texture and edge-based local features together at the descriptor extraction level. Next, we build geometric visual phrases to model spatial context upon complementary features for midlevel **image** representation. Finally, based on a smoothed edgemap, a simple and effective spatial weighting scheme is performed to capture the **image** saliency. We test the proposed framework on several benchmark data sets for **image** classification. The extensive results show the superior performance of our algorithm over the state-of-the-art methods.