Ensemble-Based Commercial Buildings Facades Photographs Classifier

Aleksei Samarin[1,3]([✉]) and Valentin Malykh[2,3]

[1] Saint-Petersburg State University, Saint Petersburg, Russia
[2] Kazan (Volga Region) Federal University, Kazan, Russia
[3] Suprema Labs, Berlin, Germany

Abstract. We present an ensemble-based method for classifying photographs containing patches with text. In particular, the proposed solution is suitable for the task of classification the images of commercial building facades by the type of provided services. Our model is based on heterogeneous ensemble usage and analysis of textual and visual features as well as special visual descriptors for areas with English text. It should be noted that our classifier demonstrates remarkable performance (0.71 in F_1 score against 0.43 baseline result). We also provide our own dataset containing 3000 images of facades with signboards in order to provide complete classification benchmark.

Keywords: Ensemble · Image classifier · Visual and textual features · Signboard image descriptor

1 Introduction

Among other problems of classifying images with text, one can single out the problem of classifying commercial buildings facades photographs by the type of provided services. That problem is of significant importance in the field of applied marketing and pattern recognition [1–4]. Unfortunately that type of an automatic signboards classification is extremely difficult due to the large number of factors such as varying shooting conditions, visual distortions, unique signage fonts and styles [1]. This variety of adverse factors leads to the need to maximize the use of all information suitable for classification. In order to implement that idea we combine several different classifiers that make decisions according to different feature types.

The first type of classifiers that we use in our ensemble is general images classifiers. At the present time, many computer vision methods have achieved outstanding results in the classification of general images [5–8]. However dissimilar objects images classification problem significantly differs from classification of photographs of facades of commercial buildings with advertising signboards by visual properties of presented objects. It should be noted that signboards are often have similar shape, that sharply distinguishes them from dissimilar objects from general datasets [9,10]. Despite this circumstance, general images