Single Image Vehicle Classification Using Pseudo Long Short-Term Memory Classifier

Reza Fuad Rachmadi, Keiichi Uchimura, Gou Koutaki, Kohichi Ogata

PII: S1047-3203(18)30236-0
DOI: https://doi.org/10.1016/j.jvcir.2018.09.021
Reference: YJVCI 2289

To appear in: *J. Vis. Commun. Image R.*

Received Date: 27 November 2017
Accepted Date: 28 September 2018

Please cite this article as: R.F. Rachmadi, K. Uchimura, G. Koutaki, K. Ogata, Single Image Vehicle Classification Using Pseudo Long Short-Term Memory Classifier, *J. Vis. Commun. Image R.* (2018), doi: https://doi.org/10.1016/j.jvcir.2018.09.021

Single Image Vehicle Classification Using Pseudo Long Short-Term Memory Classifier

Reza Fuad Rachmadi[a,b], Keiichi Uchimura[a], Gou Koutaki[a], Kohichi Ogata[a]

[a]*Graduate School of Science and Technology (GSST), Kumamoto University, Kumamoto 860-0862, Japan*
[b]*Dept. of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia*

Abstract

In this paper, we propose a pseudo long short-term memory (LSTM) classifier for single image vehicle classification. The proposed pseudo-LSTM (P-LSTM) uses spatially divided images rather than time-series images. In other words, the proposed method considers the divided images to be time-series frames. The divided images are formed by cropping input images using two-level spatial pyramid region configuration. Parallel convolutional networks are used to extract the spatial pyramid features of the divided images. To explore the correlations between the spatial pyramid features, we attached an LSTM classifier to the end of the parallel convolutional network and treated each convolutional network as an independent timestamp. Although LSTM classifiers are typically used for time-dependent data, our experiments demonstrated that they can also be used for non-time-dependent data. We attached one fully connected layer to the end of the network to compute a final classification decision. Experiments on an MIO-TCD vehicle classification dataset show that our proposed classifier produces a high evaluation score and is comparable with several other state-of-the-art methods.

Keywords: Pseudo-LSTM classifier, vehicle classification, deep convolutional network

1. Introduction

Vehicles play a significant role in modern life, and many people use them on a daily basis. The increasing number of vehicles on roads has increased the risk of traffic incidents. The Canadian Council of Motor Transport Administrators, in their 2014 statistic report [1], noted that traffic accidents in Canada cause more than 1,800 fatalities and more than 150,000 injuries. The situation forced the traffic safety division to implement a real-time surveillance system for road scene analysis. The aim of the real-time surveillance system is to understand the behavior of road users, including non-vehicle users such as cyclists and pedestrians. The combination of CCTV cameras with a video management system is the main configuration used to implement real-time traffic surveillance systems in the world today because it is cheaper and easier to implement than other sensors. Usually, traffic is manually analyzed by a control-room operator using real-time video data from CCTV cameras. To increase effectiveness, a system that can automatically analyze road traffic is required. One component of such a system is a vehicle recognition module which is very important in analyzing the behavior of road users. The vehicle recognition task is necessary to narrow down a target efficiently and effectively from a large number of vehicle categories. A vehicle recognition module usually consists of two different systems: a vehicle detection system and a vehicle classification system. The vehicle detection system is responsible for detecting the vehicle appearing in the CCTV video data, while the vehicle classification system is responsible for categorizing the vehicle. This paper discusses the vehicle classification system, by assuming that detection has already been achieved in previous steps and that the input of the vehicle classification system is the cropped region of the vehicle candidate.

The development of vehicle classification systems began in early 1980 with the use of induction loop sensors. Such sensors were used until approximately 2000, as described in [2, 3, 4, 5, 6]. One disadvantage of induction loop sensors is that they have very high implementation and maintenance costs compared with other approaches. Researchers have proposed other approaches for vehicle classification systems, including acoustic

Email address: fuad@navi.cs.kumamoto-u.ac.jp (Reza Fuad Rachmadi)