Place recognition survey: An update on deep learning approaches

Tiago Barros, Ricardo Pereira, Luís Garrote, Cristiano Premebida, Urbano J. Nunes

***Abstract*—Autonomous Vehicles (AV) are becoming more capable of navigating in complex environments with dynamic and changing conditions. A key component that enables these intelligent vehicles to overcome such conditions and become more autonomous is the sophistication of the perception and localization systems. As part of the localization system, place recognition has benefited from recent developments in other perception tasks such as place categorization or object recognition, namely with the emergence of deep learning (DL) frameworks. This paper surveys recent approaches and methods used in place recognition, particularly those based on deep learning. The contributions of this work are twofold: surveying recent sensors such as 3D LiDARs and RADARs, applied in place recognition; and categorizing the various DL-based place recognition works into supervised, unsupervised, semi-supervised, parallel, and hierarchical categories. First, this survey introduces key place recognition concepts to contextualize the reader. Then, sensor characteristics are addressed. This survey proceeds by elaborating on the various DL-based works, presenting summaries for each framework. Some lessons learned from this survey include: the importance of NetVLAD for supervised end-to-end learning; the advantages of unsupervised approaches in place recognition, namely for cross-domain applications; or the increasing tendency of recent works to seek, not only for higher performance but also for higher efficiency.**

***Index Terms*—Place recognition, Deep Learning, Localization.**

I. Introduction

Self-driving vehicles are increasingly able to deal with unstructured and dynamic environments, which is mainly due to the development of more robust long-term localization and perception systems. A critical aspect of long-term localization is to guarantee coherent mapping and bounded error over time, which is achieved by finding loops in revisited areas. Revisited places are detected in long-term localization systems by resorting to approaches such as place recognition and loop closure. Namely, place recognition is a perception based approach that recognizes previously visited places based on visual, structural, or semantic cues.

Place recognition has been the focus of much research over the last decade. The efforts of the intelligent vehicle and machine vision communities, including those devoted to place recognition, resulted in great achievements, namely evolving towards systems that achieve promising performances in appearance changing and extreme viewpoint variation conditions. Despite the recent achievements, the fundamental challenges remain unsolved, which occur when:

– two distinct places look similar (also known as perceptual aliasing);

The authors are with the University of Coimbra, Institute of Systems and Robotics, Department of Electrical and Computer Engineering, Portugal. E-mail:{tiagobarros, ricardo.pereira, garrote, cpremebida, urbano}@isr.uc.pt

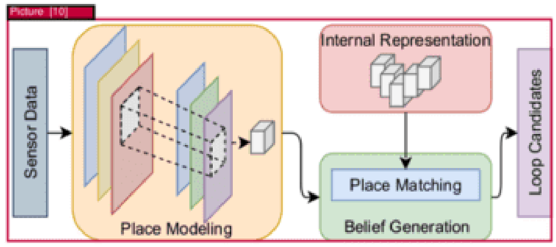

Fig. 1. Generic place recognition pipeline with the following modules: place modeling, belief generation and place mapping. Place modeling creates an internal place representation. Place mapping is concerned with maintaining a coherent representation of places over time. And Belief generation, finally generates, based on the current place model and the map, loop candidates.

Fig. 2. Illustration of the seasonal environment changes. Images taken from the Oxford Robotcar [1] and Nordland dataset [2].

– the same places exhibit significant appearance changes over time due to day-night variation, weather, seasonal or structural changes (as shown in Fig. 2);

– same places are perceived from different viewpoints or positions.

Solving these challenges is essential to enable robust place recognition and consequently long-term localization.

The primary motivation for writing this survey paper is to provide an updated review of the recent place recognition approaches and methods since the publication of previous surveys [3], [4]. The goal is, in particular, to focus on the works that are based on deep-learning.

Lowry et al. [3] presented a comprehensive overview of the existing visual place recognition methods up to 2016. The work summarizes and discusses several fundamentals to deal with appearance changing environments and viewpoint variations. However, the rapid developments in deep learning (DL) and new sensor modalities (*e.g.*, 3D LiDARs and RADARs) are setting unprecedented performances, shifting the place recognition state of the art from traditional (handcrafted-only) feature extraction towards data-driven methods.

A key advantage of these data-driven approaches is the end-to-end training, which enables to learn a task directly from the sensory data without requiring domain knowledge for fea-