Self-supervised Visual Place Recognition for Colonoscopy Sequences

Javier Morlana, Pablo Azagra, Javier Civera, José M.M. Montiel
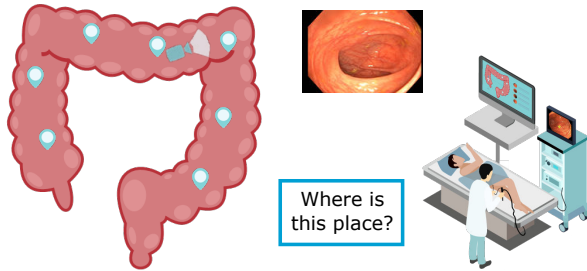
Universidad de Zaragoza

Introduction

Motivation

- Simultaneous Localization and Mapping (SLAM) will enable robotics and AR in colonoscopies
- Visual Place Recognition (VPR) allows SLAM to remove drift (loop closure) and recover from tracking failure

Where is this place?

Problem

- Input: query image from the place to be found + database of places visited
- Output: database images observing the same place as the query
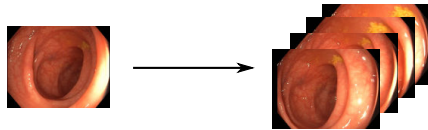
Contributions

- End-to-end netowrk for VPR in colonoscopies
- Self-supervised place labelling in colonoscopies using SfM

[1] Radenović, F., Tolias, G., & Chum, O. (2018). Fine-tuning CNN image retrieval with no human annotation. IEEE transactions on pattern analysis and machine intelligence, 41(7), 1655-1668.
[2] Schonberger, J. L., & Frahm, J. M. (2016). Structure-from-motion revisited. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4104-4113)0.

Method

Place recognition cast as image retrieval

- Images are represented by a global descriptor obtained after a forward pass through the the network proposed in [1]
- Descriptors are saved in the database
- A Nearest Neighbour search is performed for every query against the database
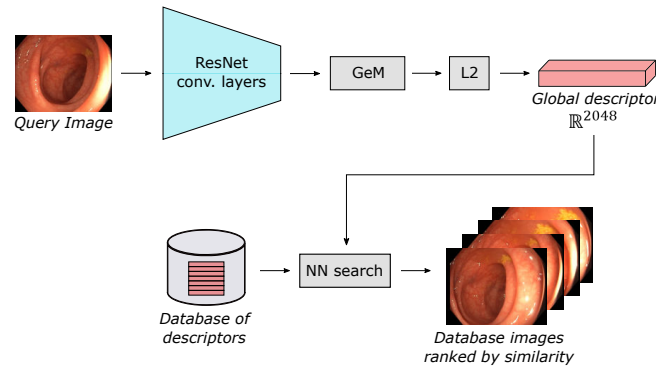
Training with contrastive loss and SfM labels

- Place labelling is self-supervised using SfM [2]
- Positives: frames from the same cluster
- Negatives: hard-mined examples from a different cluster but from the same sequence
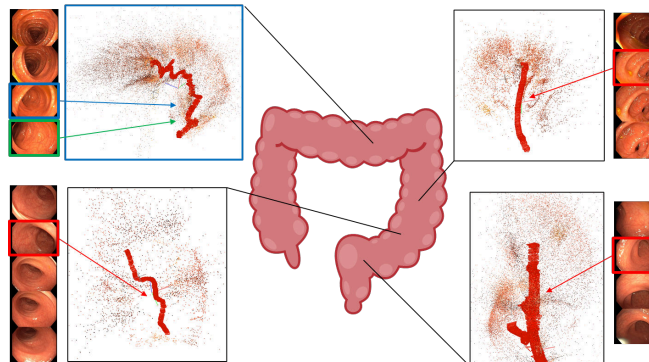

Experiments

Quantitative experiment

- Retrieval problem: any frame from the test sequence can be retrieved
- Baseline: network trained in cities [1]
- 25% (48.57-60.58) improvement in mAP
- Recall@N excluding frames closer than D
- Our system retrieves a correct first candidate (R@1) in 85.3% of the cases when D = 15

	D = 5			D = 10			D = 15		
	R@1	R@5	R@10	R@1	R@5	R@10	R@1	R@5	R@10
Radenović	92.2	96.0	97.3	83.7	90.3	92.7	77.3	85.6	88.7
Our method	96.1	98.2	98.7	90.5	94.9	96.5	85.3	91.3	93.4

Qualitative experiment

- SLAM problem: only past frames are retrieved
- Database grows as sequence is processed

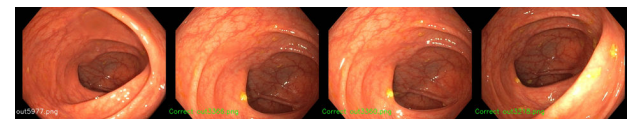- Relocalization if a place is visited for the first time
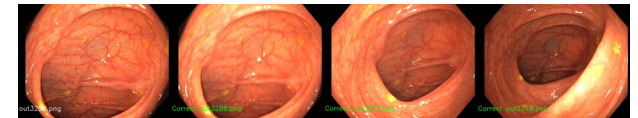
- SfM sometimes gives false negatives but we are able to retrieve them correctly
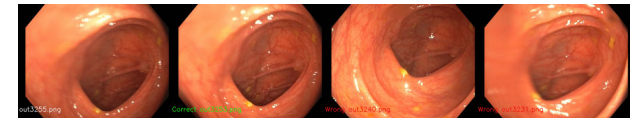
Ours / SfM   Ours / SfM   Ours / SfM

- Loop closure if a place is revisited

D=2611   D=2617   D=2759