The Super High Intensity Team

Terran Mott & Robinson Merillat

Machine Learning Theory Project Proposal

**Paper Description**

**Paper Title:** Learning Scene Geometry for Visual Localization in Challenging Conditions

**Authors:** Nathan Piasco, Désiré Sidibé, Valérie Gouet-Brunet, Cedric Demonceaux

**Conference:** ICRA 2019

**Description:**

This paper proposes a method of visual-based localization that is robust to changes in weather, lighting, and seasonality in an environment. In this case, VBL is framed as an image-retrieval problem, in which an input image is compared to a pool of known references. The value and novelty of this approach is that it would 1) be available to robots without the capability to generate 3D representations of their surroundings and 2) be available in cases where images’ semantic representations cannot be obtained.

VBL is important in a vast array of applications involving computer vision problems in changing environments. An autonomous vehicle or a robot navigating an outdoor setting must be able to recognize images and objects in spite of changes in lighting and weather. For example, it is a particularly difficult problem to build reliable image recognition techniques which are robust against the presence of snow, or at nighttime.

The particular algorithm developed in this paper uses auxiliary depth information to improve performance. This depth information is only available during training and is used to build descriptors of an image. The final algorithm takes in a query image and responds by ranking its reference images according to a similarity score. In this way, this technique deals only in images, instead of 3D representations which are not always available.

**Dataset:**

The algorithm in this paper was both trained and tested on the Oxford Robotcar dataset, which is commonly used for VBL. The data consist of images of the same street views taken in different weather conditions and times of day.

<https://robotcar-dataset.robots.ox.ac.uk/>

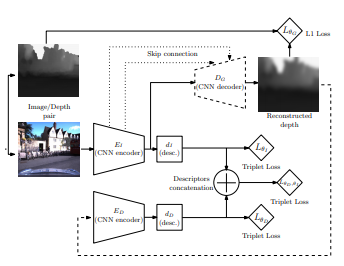
**Experiment:**

The experiment used in this paper to evaluate their algorithmic approach was to train their novel image descriptor and run it on a set of oxford robotcar data. The goal was to accurately identify the location of the image (ground truth that is available through Oxford) using the descriptor they made.

They evaluated their descriptor against several other state of the art localization methods, including a hallucination approach. Recall and top-1 recall were used to compare the performance. They found that recall performance metrics increased by ~2% across weather variations and by ~4% on seasonal variations. Surprisingly, they also found a significant increase across night/day variance.

We plan to recreate this paper’s algorithmic approach using a series of convolutional neural networks to complete a supervised localization task. Our first task will be to select a subset of the Oxford Robotcar data (which is quite extensive) that we would like to use. Once our data is selected and cleaned, we can begin to construct the algorithmic components of the network, including the depth-map constructor. We can use resources already available online to build our depth maps, which are only one component of the network.

The proposed descriptor will be constructed as in the the following flowchart



**Timeline**:

**Oct 20:** Have data selected and a plan for cleaning

- For the Oxford Robotcar data, this means selecting a given subset of images

**Oct 26:** Theory Project Update

- have a cleaned dataset of images

- in-progress algorithm implementation: Have depth-map capability for our training dataset

**Mid-November:**

- test implementation and evaluate model