

# 1 Related Work

The learning and perception of traversability is a fundamental competence for both organisms and autonomous mobile robots since most of their actions depend on their mobility [ugur2010traversability]. Visual perception is known to be used in most all animals to correctly estimate if an environment can be traversed or not. Similar, a wide array of autonomous robots adopt local sensors to mimic the visual properties of animals to extract geometric information of the surrounding and plan a safe path through it. Different methodologies have been proposed to collect the data and then learn to correctly navigate the environment. Most of the methodologies rely on supervised learning, where first the data is gathered and then a machine learning algorithm is trained sample to correctly predict the traversability of those samples.

Among the huge numbers of methods proposed, there are two categories based on the input data: geometric and appearance based methods.

Geometric methods aim to detect traversability using geometric properties of surfaces such as distances in space and shapes. Those properties are usually slopes, bumps, and ramps. Since nearly the entire world has been surveyed at 1 m accuracy [sofman2006improving], outdoor robot navigation can benefit from the availability of overhead imagery. For this reason, elevation data has also been used to extract geometric information. Recently, [1709.05368], the work we base on, proposed a full pipeline to estimate traversability using only elevation data in the form of height maps.

Elevation data can also be estimated by flying drones. [delmerico2016active] proposed a collaborative search and rescue system in which a flying robot that explores the map and creates an elevation map to guide the ground robot to the goal. They utilize an 'on-the-spot training' using a convolutional neural network to segment the terrain in different traversable classes.

Whereas appearance methods, to a greater extent related to camera images processing and cognitive analyses, have the objective of recognizing colors and patterns not related to the common appearance of terrains, such as grass, rocks or vegetation. Those images can be used to directly estimate the traversability cost.

Historically, the collected data is first preprocessed to extract texture features that are used to fit a classic machine learning classified such as an SVM [ugur2010traversability] or Gaussian models [sofman2006improving]. Those techniques rely on texture descriptors, for example, Local Binary Pattern [ojala2002multiresolution], to extract features from the raw data images obtained from local sensors such as cameras. With the rise of deep learning methods in computer vision, deep convolution neural network has been trained directly on the raw RGB images bypassing the need to define characteristic features.

One recent example is [giusti2015amachine] where a deep neural network was training on real-world hiking data collected using head-mounted cameras to teach a flying drone to follow a trail in the forest. Geometric and appearance methods can also be used together to train a traversability classifier. [delmerico2017onthespot] proposes the first on-the-spot training method

that uses a flying drone to gather the data and train an estimator in less than 60 seconds.

Data can also be extracted in simulations, where an agent interacts in an artificial environment. Usually, no-wheel legged robot able to traverse harder ground, can benefit from data gathering in simulations due to the high availability. For example, [tobias2017anytime] proposed a locomotion planning that is learned in a simulation environment.

Legged robots show their full potential in rough and unstructured terrain, where they can use a superior moves set compared to wheel robots. Different frameworks have been proposed to compute safe and efficient paths for legged robots. [wermelinger2016navigation] uses typical map characteristics such as slopes and roughness gathered using onboard sensors to train a planner. The planner uses a RRT\* algorithm to compute the correct path for the robot on the fly. Moreover, the algorithm is able to first find an easy local solution and then update its path to take into account more difficult scenarios as new environment data is collected.

While the previous method relies on features extraction from an heightmap to estimate the cost of a given patch using a specific function, new framework that automatise the features extraction process has been proposed recently. [wellhausen2019where] they use local sensing to train a deep convolutional neural network to predict terrain's properties. They collect data from robot ground interaction to label each image in front of the robot in order to predict the future interactions with the terrain showing that the network is perfectly able to learn the correct features for different terrains. Furthermore, they also perform weakly supervised semantic segmentation to divide the input images in different ground classes, such as glass and sand, showing respectable results.