# Single and Multi-Label Environmental Sound Classification Using Convolutional Neural Networks

SANTIAGO ÁLVAREZ-BUYLLA PUENTE
Department of Civil and Environmental Engineering
Chalmers University of Technology

## Abstract

Artificial neural networks are computational systems made up of simple processing units that have a natural propensity for storing experiential knowledge and making it available for use. In the recent years this technology has seen an exponential growth in the fields of image recognition, natural language processing or speech recognition. However, there is a dearth of research on environmental sound analysis. In combination with IoT and wireless sensor networks, this technology would help to identify and therefore better address noise issues present in urban environments. This master thesis investigates the theory and construction of artificial neural networks for single-label and multi-label multiclass classification of environmental sounds like dog bark, street music or jackhammer. Evaluation to different corruptions of the sounds are studied, as well as methods to increase robustness to these variations.

A convolutional neural network arquitecture is proposed for both tasks. The inputs to the networks are time-frequency patches extracted from the computed mel-spectrogram of the signals. Dropout and weight decay regularization methods are applied and the cross-entropy loss is optimized using Adam algorithm.

Results show that these systems are very sensitive to noise and level corruptions of the inputs. Data augmentation techniques apply techniques like data augmentation to avoid these issues. Results to the multi-label classification task show that it is still possible for a neural network to learn in a complicated mixed environment. However there is still room for improvement regarding prediction accuracy. Since no previous benchmarks are available for comparison, this study sets the stage for the multi-label classification task using UrbanSound8K dataset.