REGULARIZED CNN FOR SOUND EVENT DETECTION IN DOMESTIC ENVIRONMENTS

Technical Report

Alexander Rakowski

Samsung R&D Poland, Audio Intelligence Dept., Warsaw, Poland, a.rakowski2@samsung.com

ABSTRACT

This report describes a system used for Task 4 of the DCASE 2019 Challenge - Sound Event Detection in Domestic Environments. The system consists of a 9-layer convolutional neural network which yields frame-level predictions. These are then aggregated using a Voice Activity Detection algorithm in order to extract sound events. To prevent the system from overfitting two techniques are applied. The first one consists of training the model with channel- and pixel-wise dropout. The second one removes information from a randomly selected subset of frames.

Index Terms— Convolutional neural network, dropout, spectrogram occlusions