

Accepted Manuscript

Noise Robust Sound Event Classification with Convolutional Neural Network

Ilyas Ozer, Zeynep Ozer, Oguz Findik

PII: S0925-2312(17)31267-5
DOI: [10.1016/j.neucom.2017.07.021](https://doi.org/10.1016/j.neucom.2017.07.021)
Reference: NEUCOM 18709

To appear in: *Neurocomputing*

Received date: 13 March 2017
Revised date: 28 June 2017
Accepted date: 11 July 2017

Please cite this article as: Ilyas Ozer, Zeynep Ozer, Oguz Findik, Noise Robust Sound Event Classification with Convolutional Neural Network, *Neurocomputing* (2017), doi: [10.1016/j.neucom.2017.07.021](https://doi.org/10.1016/j.neucom.2017.07.021)



This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Noise Robust Sound Event Classification with Convolutional Neural Network

Ilyas Ozer, Zeynep Ozer, Oguz Findik

Karabuk University, Computer Engineering Department, Karabuk, Turkey

Abstract

Automatic sound recognition (ASR) is a remarkable field of research in recent years. The ability to automatically recognize sound events through computers in a complex audio environment is very useful for machine hearing, acoustic surveillance and multimedia retrieval applications. On the other hand, ASR task become highly difficult as the ambient noise levels increase and many traditional methods show very weak performance under noise. Recent studies has shown that spectrogram image features (SIF) have high performance under noise, while success rates in clean conditions are relatively lower than in the state-of-the-art approaches. In this study, after converting highly overlapped spectrograms into linear quantized images and reducing dimensions by applying various image resizing methods, feature extraction and classification are performed with convolutional neural networks (CNN), which have very high performance in image classification. In the mismatched case, the proposed method achieves a performance improvement of 4.5%, which is equivalent to a relative error reduction of 63.4%, with a classification success of 97.4%, while the multicondition training method achieves an average of 98.63% success rate.

Keywords: Sound event classification, Convolutional neural networks, Spectrogram

1. Introduction

While sound was divided into two categories as speech and non-speech in initial studies in the literature, there are much more studies on speech recognition between these two interrelated subjects. Research on non-speech sound recognition, also known as ASR, has gained momentum in recent years

Download English Version:

<https://daneshyari.com/en/article/6865324>

Download Persian Version:

<https://daneshyari.com/article/6865324>

[Daneshyari.com](https://daneshyari.com)