IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 8, No. 3, September 2019, pp. 264~269
ISSN: 2252-8938, DOI: 10.11591/ijai.v8.i3.pp264-269

Review of anomalous sound event detection approaches

Amirul Sadikin Md Afendi[1], Marina Yusoff[2]
[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia
[2]Advanced Analytic Engineering Center (AAEC), Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

Article Info

Article history:

Received Jun 27, 2019
Revised Aug 5, 2019
Accepted Aug 20, 2019

Keywords:

Acoustic Features
Audio Processing
Pattern Recognition
Sound Event Detection

ABSTRACT

This paper presents a review of anomalous sound event detection (SED) approaches. SED is becoming more applicable for real-world appliactaions such as security, fire determination or olther emergency alarms. Despite many research outcome previously, further research is required to reduce false positives and improve accurracy. SED approaches are comprehensively organized by methods covering system pipeline components of acoustic descriptors, classification engine, and decision finalization method. The review compares multiple approaches that is applied on a specific dataset. Security relies on anomalous events in order to prevent it one must find these anomalous events. Audio surveillance has become more efficient as that artificial intelligence has stepped up the game. Autonomous SED could be used for early detection and prevention. It is found that the state of the art method viable used in SED using features of log-mel energies in convolutional recurrent neural network (CRNN) with long short term memory (LSTM) with a verification step of thresholding has obtained 93.1% F1 score and 0.1307 ER. It is found that feature extraction of log mel energies are highly reliable method showing promising results on multiple experiments.

Corresponding Author:

Marina Yusoff,
Advanced Analytic Engineering Center (AAEC),
Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia.
Email: marinay@tmsk.uitm.edu.my

1. INTRODUCTION

The Sound event detection (SED) research field has been an active recently [1-4]. Autonomous audio surveillance has become more efficient as that artificial intelligence has stepped up the game [5]. Machine learning has become very powerful to allow new approaches to be facilitated in countless domain [6]. Acoustic surveillance specifically in security application is still new to the world and requires research to allow better performances [4, 7, 8]. The detection on anomalous audio events effectively while avoiding false positives is crucial in security [3]. Two categories of sound recognition, non-speech the determination of the sound event source and speech recognition of verbal language [9-10]. Many works embrace the method of multi-label classification to detect polyphonic acoustic events with a worldwide limit to detect active acoustic events [2, 7, 10]. The lack of accuracy and high false positives on current SED approaches is holding autonomous audio surveillance to be applied for real-world applications [11]. In the race towards making audio surveillance a reality many have aid by providing data for training and testing new models in solving the problem. Datasets on SED can be found available in many varieties as the internet of things are becoming more popular [1, 10, 12].

The rare term relates to the possibility that any target sound event to be identified may happen at most once within half a minute [4, 12]. Detecting rare sound events are one of the key features