

# MIREX 2017 DRUM TRANSCRIPTION SUBMISSIONS

Carl Southall

DMT Lab, Birmingham City University  
Birmingham, United Kingdom  
carl.southall@bcu.ac.uk

## ABSTRACT

This abstract outlines the submissions made for the drum transcription MIREX 2017 task.

## 1. SUBMISSIONS

The three submissions all follow the same automatic drum transcription (ADT) system outlined used in [2,3]. All systems are trained using the Adam optimizer with an initial learning rate of 0.003. 75% of training tracks are used in mini-batches, consisting of ten 100-frame segments, and the other 25% are used for optimisation of the peak-picking parameters. An open source web based version of the systems is available (ADTWeb) [1] as well as a python library.<sup>1</sup>

### 1.1 CS1

CS1 is based on the SA system from [3]. The output layer is connected to  $a$  time-step hidden layers on either side of the current frame ( $a=3$ ). Target activation functions used in training are created by setting a single frame to 1 ( $Y^t=1$ ) for each onset and the mean threshold (MT) peak-picking technique is used.

### 1.2 CS2

CS2 follows the same system design as CS1, however target activation functions used in training are created by setting a total of three frames to 1 ( $Y^{t-1:t+1}=1$ ) for each onset.

### 1.3 CS3

CS3 follows the same system design as CS2, however there is a 50% chance of data augmentation occurring in each mini-batch before training. The data is augmented by means of pitch shifting and the addition of white noise.

## 2. RESULTS

Table 1 presents the results achieved by the submissions for both the public and evaluation subsets. The CS2 system demonstrates the best overall performance of the three submissions and the best in the whole evaluation for the

**Table 1.** Public and evaluation set results from the three submissions. Bold represents the highest performance in the whole evaluation.

System	Global	IDMT	KT	Gen	MDB	RBMA
Public (Training) Set						
CS1	0.71	0.68	0.68	0.70	0.76	0.68
CS2	0.77	0.77	0.77	<b>0.78</b>	0.81	0.72
CS3	0.71	0.71	0.71	0.71	0.78	0.64
Evaluation (Test) Set						
CS1	0.61	0.51	0.53	0.68	0.68	0.66
CS2	0.63	0.51	0.56	0.69	<b>0.74</b>	0.66
CS3	0.63	0.52	0.55	0.72	0.72	0.64

MDB Drums dataset. The results also suggest that the systems may have overfit, possibly due to the lack of training data, which is an open problem in ADT research [4].

## 3. REFERENCES

- [1] Carl Southall, Nick Jillings, Ryan Stables, and Jason Hockman. ADTWeb - An open source browser based automatic drum transcription system. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, Suzhou, China, 2017.
- [2] Carl Southall, Ryan Stables, and Jason Hockman. Automatic drum transcription using bi-directional recurrent neural networks. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 591–597, New York City, United States, 2016.
- [3] Carl Southall, Ryan Stables, and Jason Hockman. Automatic drum transcription for polyphonic recordings using soft attention mechanisms and convolutional neural networks. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 606–612, Suzhou, China, 2017.
- [4] Carl Southall, Chih-Wei Wu, Alexander Lerch, and Jason Hockman. MDB Drums: An annotated subset of MedleyDB for automatic drum transcription. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, Suzhou, China, 2017.

<sup>1</sup> <https://github.com/CarlSouthall/ADTLib>

