MIREX 2026 Submission: Context-Aware Semi-Markov Post-Processor for Beat and Downbeat Estimation

Zhanhong He[1], Hanyu Meng[2], Jiahao Zhao[3], Ganghui Ru[4], Yaolong Ju[5]

[1] University of Western Australia [2] University of New South Wales [3] Kyoto University [4] Fudan University [5] Great Bay University

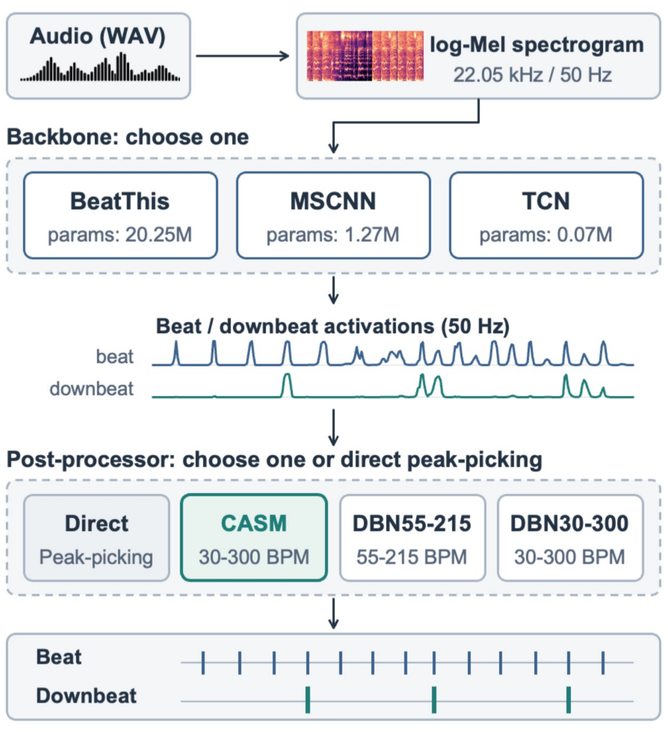
Submitted System Overview

Neural Beat Tracker (Backbones): BeatThis (20.25M), MSCNN (1.27M), TCN (0.07M parameters).

Decoding Methods: Direct (minimal peak-picking), CASM Post-Processor, and DBN Post-Processor under 55-215 bpm or 30-300 bpm setup.

"Context-Aware Semi-Markov is proposed in our under-review paper. Detailed methodologies in there."

Core Concept: DBNs rely on hard-coded predefined priors (beat_per_bar, bpm ranges, ...) for optimal. One setting is unstable across data, so necessitating an adaptive method >> PLPDP >> we proposed CASM.



MIREX 2026 Results

Method	Beat / Downbeat (GTZAN)						
	F1	Cemgil	CMLt	AMLt	F1	Prec	Recall
BeatThis	88.7	79.9	79.5	89.4	**77.3**	75.3	**82.4**
w. CASM	**+0.0**	**+0.0**	+0.2	+0.4	−0.1	+2.2	−2.9
w. DBN[55,215]	−0.6	−0.8	**+1.2**	+2.1	−0.8	+3.1	−5.7
w. DBN[30,300]	−0.5	−0.7	+0.8	**+2.6**	−0.5	**+3.4**	−5.3
MSCNN	86.2	78.0	72.6	79.3	68.7	63.4	**77.9**
w. CASM	**+0.3**	**+0.3**	+1.0	+1.4	+2.3	+6.0	−2.6
w. DBN[55,215]	−0.7	−1.6	**+4.3**	**+7.2**	**+2.9**	**+10.0**	−6.5
w. DBN[30,300]	−1.7	−2.3	+2.3	+6.2	+2.2	+9.8	−7.2
TCN	86.3	76.0	72.5	80.3	60.8	49.6	**85.3**
w. CASM	**+0.3**	**+0.2**	**+1.0**	+1.2	+4.6	+12.3	−9.0
w. DBN[55,215]	−2.2	−2.9	+0.2	**+5.9**	**+5.7**	**+17.9**	−17.4
w. DBN[30,300]	−2.8	−3.4	−1.4	+5.6	+5.1	+17.4	−18.1

Method	Beat (SMC)					
	F1	Cemgil	Goto	P-score	CMLt	AMLt
BeatThis	57.7	44.4	15.7	65.1	41.8	52.3
w. CASM	**+0.2**	**+0.2**	+4.2	**+0.2**	+1.5	+2.5
w. DBN[55,215]	−3.1	−2.6	+8.3	−1.4	+0.3	+5.8
w. DBN[30,300]	−3.0	−2.5	**+10.1**	−2.0	**+1.7**	**+5.9**
MSCNN	54.8	42.3	5.1	60.7	34.9	41.6
w. CASM	**+0.9**	**+0.7**	+3.7	**+1.8**	**+4.4**	+5.2
w. DBN[55,215]	−4.3	−4.1	**+15.7**	+0.2	+4.3	**+11.9**
w. DBN[30,300]	−7.9	−6.7	+13.8	−5.9	−1.8	+8.0
TCN	51.5	40.0	4.6	57.8	27.8	33.4
w. CASM	**+0.3**	**+0.2**	+3.2	**+0.4**	**+2.2**	+4.4
w. DBN[55,215]	−9.6	−7.8	**+8.8**	−7.9	−5.8	**+6.1**
w. DBN[30,300]	−14.9	−11.9	+6.0	−15.8	−12.0	+3.4

Method	Beat / Downbeat (osu2beat2025)						
	F1	Cemgil	CMLt	AMLt	F1	Prec	Recall
BeatThis	**88.2**	**59.9**	**70.0**	76.5	**83.8**	86.9	**82.3**
w. CASM	−0.3	−0.2	−0.7	0.0	−1.8	+0.7	−3.9
w. DBN[55,215]	−2.1	−1.5	−3.8	**+2.8**	−1.6	**+3.1**	−4.8
w. DBN[30,300]	−1.9	−1.4	−3.5	+2.7	−1.4	**+3.1**	−4.5
MSCNN	80.3	**55.8**	**53.7**	58.5	**72.5**	71.4	**75.1**
w. CASM	**+0.1**	−0.1	−0.0	+0.2	−0.6	+2.1	−3.3
w. DBN[55,215]	−4.5	−3.3	−0.3	**+9.3**	−1.5	+9.8	−9.3
w. DBN[30,300]	−5.1	−3.7	−0.8	+8.4	−1.9	**+10.1**	−9.9
TCN	84.4	67.5	61.0	64.5	60.9	48.9	**85.2**
w. CASM	**+0.1**	**+0.1**	**+0.1**	+0.1	+3.0	+7.9	−4.4
w. DBN[55,215]	−5.5	−4.5	−8.2	**+12.2**	**+6.2**	**+27.0**	−22.5
w. DBN[30,300]	−6.1	−5.0	−8.2	+10.4	+5.6	+26.6	−22.9

Method	Beat / Downbeat (JSIR)						
	F1	Cemgil	CMLt	AMLt	F1	Prec	Recall
BeatThis	**92.8**	**85.9**	**79.0**	85.0	84.3	81.1	**88.5**
w. CASM	−0.3	−0.3	−0.9	+0.2	**+1.3**	+7.6	−4.9
w. DBN[55,215]	−9.1	−9.0	−22.2	−1.0	−6.6	+8.7	−16.8
w. DBN[30,300]	−4.2	−4.4	−9.8	**+0.7**	−1.3	**+9.8**	−9.9
MSCNN	89.3	79.2	65.3	69.0	69.4	61.2	**81.7**
w. CASM	**+0.3**	**+0.4**	**+1.2**	**+1.7**	**+6.6**	+11.4	−0.0
w. DBN[55,215]	−11.6	−11.2	−15.9	−0.5	+0.8	+22.4	−17.2
w. DBN[30,300]	−10.7	−10.5	−13.5	+0.3	+1.9	**+22.6**	−15.8
TCN	86.6	76.7	59.7	63.4	49.0	35.5	**83.3**
w. CASM	**+0.6**	**+0.6**	**+2.3**	+2.2	+5.5	+8.8	−1.2
w. DBN[55,215]	−14.1	−14.0	−22.6	+2.3	**+5.6**	**+24.8**	−28.5
w. DBN[30,300]	−13.7	−13.7	−22.4	**+3.6**	+5.2	+24.3	−28.8

Method	Beat / Downbeat (Yamaha Balanced)						
	F1	Cemgil	CMLt	AMLt	F1	Prec	Recall
BeatThis	90.0	77.6	84.1	87.5	85.3	82.3	**89.9**
w. CASM	+0.4	**+0.3**	+1.1	+1.3	+0.3	+2.9	−2.9
w. DBN[55,215]	+0.5	+0.1	**+3.1**	**+3.5**	+1.7	+4.5	−1.8
w. DBN[30,300]	**+0.6**	+0.2	+2.9	+3.1	**+1.8**	**+4.7**	−1.8
MSCNN	87.5	73.7	77.7	79.8	78.3	72.7	**86.0**
w. CASM	**+0.7**	**+0.6**	+2.1	+2.1	+2.9	+7.2	−2.5
w. DBN[55,215]	+0.1	−0.8	**+5.3**	**+6.4**	**+5.4**	+11.3	−1.7
w. DBN[30,300]	−0.2	−1.1	+5.3	+6.0	+5.3	**+11.7**	−2.2
TCN	87.3	76.9	75.1	79.6	63.6	50.9	**89.5**
w. CASM	**+0.4**	**+0.4**	**+1.6**	+1.7	+9.1	+16.7	−4.2
w. DBN[55,215]	−2.6	−3.1	−3.0	**+7.0**	**+12.7**	**+23.3**	−8.5
w. DBN[30,300]	−3.1	−3.5	−3.8	+6.5	+12.1	+22.7	−9.0

Method	Beat / Downbeat (Yamaha JPop)						
	F1	Cemgil	CMLt	AMLt	F1	Prec	Recall
BeatThis	94.2	82.8	87.8	89.2	91.5	90.0	**94.1**
w. CASM	**+0.2**	**+0.2**	+0.7	+0.7	+0.2	+2.2	−2.0
w. DBN[55,215]	+0.1	−0.1	**+1.4**	+3.0	**+1.1**	**+3.2**	−1.2
w. DBN[30,300]	+0.1	−0.1	+1.3	**+3.0**	+1.1	+3.1	−1.2
MSCNN	90.1	75.9	77.2	78.4	83.5	78.6	**90.4**
w. CASM	+0.8	**+0.7**	+1.8	+1.9	+2.2	+6.3	−2.7
w. DBN[55,215]	**+0.9**	−0.4	**+6.8**	+9.5	+4.9	+11.1	−1.9
w. DBN[30,300]	+0.9	−0.5	+6.8	**+9.5**	**+5.0**	**+11.4**	−2.0
TCN	91.3	82.4	79.4	82.1	70.1	57.2	**94.5**
w. CASM	**+0.4**	**+0.4**	**+1.6**	+1.6	+9.4	+18.2	−5.4
w. DBN[55,215]	−2.5	−3.2	−2.5	**+6.6**	**+13.2**	**+25.3**	−8.0
w. DBN[30,300]	−3.2	−3.8	−3.9	+6.5	+12.7	+24.6	−8.0

Conclusion

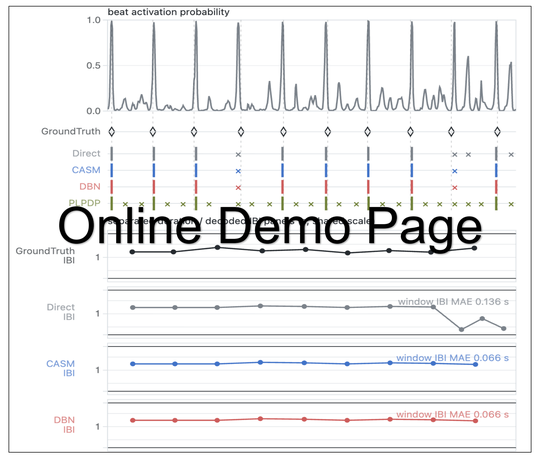
CASM is effective on 16/18 backbones-dataset pairs.

(1) CASM generally improves on the backbone model's performance, which is a consistently improvement (good balance between the event accuracy (F1) and consistency (CMLt, AMLt, …)

(2) Only 2/18 special cases in osu2beat2025 and JSIR datasets, where both CASM and DBN post-processed results are worse than the backbones' direct output.

Methodology is in our under-review paper. GitHub live demo is available, with pip-installation package & pretrained backbones provided: https://zhanh-he.github.io/casm-beat-tracking/

- Thanks to MIREX organizers for these thorough evaluation -



Contact: zhanh.he.uwa@gmail.com, juyaolong@gbu.edu.cn