HiTaiko: Hierarchical osu!taiko Chart Generation via Metrical Candidate Selection

Sungjung Park
Department of Applied Statistics, Chung-Ang University, Seoul, Republic of Korea

ISMIR 2026
Abu Dhabi, UAE

1. Dataset

clean269: five-difficulty osu!taiko corpus

We use a curated corpus (clean269) containing 269 osu!taiko beatmapsets, where exactly one chart from each of the five standard difficulty tiers (T1–T5) is selected per song. The five charts of each song share an identical uninherited timing map.

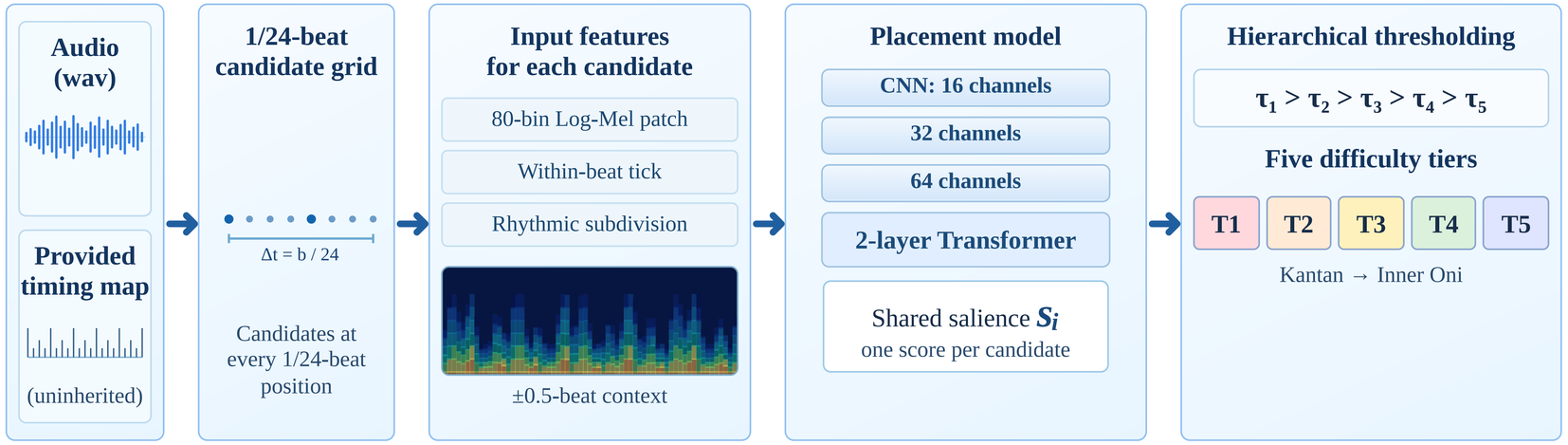
269	1,345	733,581
songs (beatmapsets)	charts (5 per song)	tap notes (all charts)

90.84% of positive metrical positions follow five-tier monotonic patterns **(T1 ⊆ T2 ⊆ T3 ⊆ T4 ⊆ T5)**

Adjacent-tier note retention rate

T1 → T2	T2 → T3	T3 → T4	T4 → T5
93.41%	94.85%	95.72%	96.05%

2. System Architecture



3. Hierarchical Difficulty Generation

A shared placement salience s_i is converted into tier-specific placement probabilities using learned ordered thresholds.

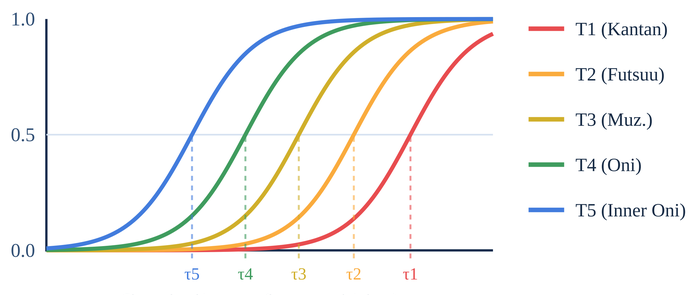
$$p_{i,t} = \sigma(s_i - \tau_t)$$

σ(·): sigmoid function

$$\tau_1 > \tau_2 > \tau_3 > \tau_4 > \tau_5$$

T1 ⊆ T2 ⊆ T3 ⊆ T4 ⊆ T5

Hierarchical thresholding on shared salience



T1 (Kantan) has the strictest threshold; T5 (Inner Oni) includes all positions retained in lower tiers.

4. Official MIREX 2026 Results (Expert Evaluation)

Human-Likeness (across-song normalized scores)

Song	HiTaiko	DOSU	FumenLab	MALer
Silver Bullet	**3.01**	−0.47	−1.61	−1.44
A fight in Tunisia	**1.70**	−2.38	−0.86	−1.28
YUBIKIRI-GENMAN	**2.81**	−1.61	0.79	−2.61
Something Fun	0.49	−3.51	1.62	**2.18**
closer to you	1.83	0.37	**4.14**	−3.18
Mean (5 songs)	**1.97**	**−1.52**	**0.82**	**−1.27**

🏆 **Top Human-Likeness on 3 of 5 songs**
Highest five-song mean among 4 systems (1.97)

Expert ratings: per-judge normalized scores, summed across three judges per song. The five-song mean is a descriptive average, not an official overall score.

5. Community Evaluation / Limitations

Himitsu feat. KAFU (Inner Oni, n = 5)

Average ratings from 5 community evaluators (1–10 scale, higher is better).

System	Overall Quality	Human-Likeness
HiTaiko	**3.33**	**2.40**
DOSU	3.67	1.00
FumenLab	5.20	6.00
MALer	5.67	3.00
Human mapper	**8.80**	**9.80**

Limitation (from community evaluation)

On this test song, HiTaiko scored below the human-authored chart (3.33 vs 8.80 overall; 2.40 vs 9.80 Human-Likeness). Only five Inner Oni participants were surveyed.

Source: MIREX 2026 Rhythm Game Chart Generation Results. Human-Likeness is not included in Overall Quality.

Shared salience + learned ordered thresholds provide nested tap placements across five tiers.