K10 A Compact Convolutional Network for Global Key Estimation

John Hurliman
Independent researcher
Oakland, CA, USA
jhurliman@jhurliman.org

MIREX 2026 · Audio Key Detection

K10 is part of JAMS. Code and bundled K10, TP1 and D1 weights are MIT licensed; third-party components retain their own licenses.

A 109,128-parameter CNN (~450 kB weights) estimates one global major or minor key from a full-track log-CQT.
Its dense readout retains frequency position after averaging over time.
One whole-track CPU forward pass; no ensemble. Memory grows with track length.

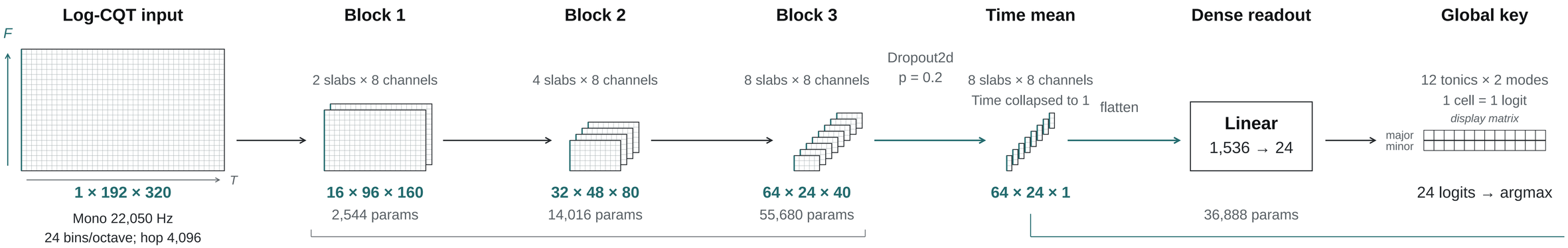
Network drawing key
Spatial cell = 8 frequency × 8 time positions per channel, at that stage.
Feature-map slab = 8 channels; input plane = 1 channel.
Time-mean and output cell units are specified below.

Architecture

Three convolutional blocks downsample both axes; the final mean averages over time only.

Channels × frequency × time

320-frame training-crop example. Inference uses the full track, with variable T.



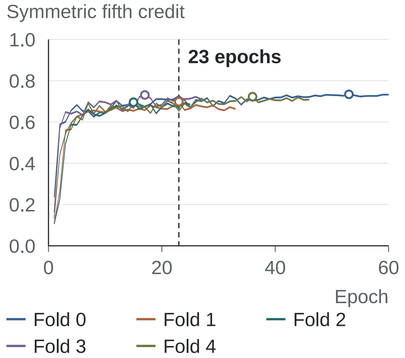
jams annotator



Training and selection

1,363 Beatport excerpts; five folds, not artist-grouped. Training: 320-frame crops, ±4-semitone shifts.

Held-out weighted score
Symmetric fifth credit



Fold-best epochs: 53, 23, 15, 17, 36.
Median 23 sets the final fit on all 1,363 tracks. Fold 0 reached the 60-epoch cap.

Official MIREX evaluation

Organizer scores for Hurliman_K10; all values are percentages.

Dataset	n	Weighted sym.	Weighted asym.	Exact
GiantSteps Key	604	**74.57**	72.91	67.88
GTZAN	837	**72.03**	70.24	64.64
FMAKv2	5,489	**72.43**	70.33	64.40
JSIR	226	**34.78**	33.67	21.68
Yamaha Balanced	185	**89.30**	89.30	85.95
Yamaha JPop	174	**85.80**	85.80	79.89

Weighted symmetric: 0 25 50 75 100

Fifth credit is 0.5 in both directions (sym.) or ascending only (asym.);
exact = 1, relative = 0.3, parallel = 0.2. Official GiantSteps uses 604 excerpts, not the paper's separate 567-excerpt development subset.
JSIR is the weakest official result. Aggregate scores do not identify a musical cause.

JAMS beyond key estimation

Key and tempo estimation; beats, downbeats and functional song sections; four-stem separation with per-stem and combined MIDI transcription.

Data-use disclosure

Earlier development used GiantSteps Key, including training a retired model. K10 training and selection used Beatport. Prior use was disclosed; the paper reports ranked eligibility confirmed on 28 Sep 2026.

Annotation workflow

Import audio → analyze → correct beats and section labels/boundaries → save annotations. Per-stem piano rolls are read-only.