Localization of gravitational waves using machine learning

Seiya Sasaoka[A], Yilun Hou[A], Kentaro Somiya[A], and Hirotaka Takahashi[B]
Tokyo Institute of Technology[A] and Tokyo City University[B]

Low-latency detection and sky localization of gravitational waves are important for electromagnetic follow-up observations. The current data-analysis method relies on matched filtering, and the computational cost is often a problem. As an alternative, machine learning is increasingly being applied in the analysis of various gravitational-wave data. In this study, we used machine-learning method for sky localization of gravitational-wave signals from binary black hole mergers using four detectors: LIGO H1, LIGO L1, Virgo, and KAGRA. Our method consists of two deep learning models: artificial neural network (ANN) and temporal convolutional network (TCN). In this poster, we report and discuss the results.

DATA GENERATION

We used ggwd package to generate datasets. The data generation process is as follows:

1) simulate waveforms using SEOBNRv4 model and parameters in Table I with a sampling rate of 2048 Hz
2) simulate noise using the design sensitivity of each detector and inject signal
3) whiten and lowpass at 500 Hz
4) cut to a length of 0.25 secs: 0.2 secs before the merger and 0.05 secs after the merger.

TABLE I. Parameters used to simulate the signals.

Mass1, Mass2	[$30M_\odot$, $80M_\odot$]
Spin1z, Spin2z	[0, 0.998]
Right ascension	[0, 2π]
Declination	[$-\pi/2$, $\pi/2$]
Coalescence phase	[0, 2π]
Inclination	[0, π]
Polarization	[0, 2π]
Network SNR	[10, 50]

We generated 800,000 samples for training, 100,000 samples for validation, and 100,000 samples for testing.

METHODS

To treat the problem as a classification task, we divided the sky into 3072 pixels using HEALPix; the solid angle of one pixel is 13.6 deg^2. We used the following three methods to classify GW signals into one of the pixels.

- Method I: ANN

This method was proposed by C. Chatterjee *et al.* [1]. Input of the ANN (Fig. 1a) are the following seven features:

1. arrival time delays of signals
2. maximum cross-correlation values of signals
3. arrival time delays of analytic signals
4. maximum cross-correlation values of analytic signals
5. ratios of amplitudes around merger
6. phase lags around the merger
7. complex correlation coefficients of signals

- Method II: TCN

For method II, the raw time-series data were input to the TCN. A TCN is a one-dimensional convolutional neural network model which consists of several temporal blocks (Fig. 1b). We stacked seven temporal blocks with dilation 2^i and kernel size 3 for the i-th block (see Fig. 1c). The input had three (or four) channels when using strain data from three (or four) detectors.

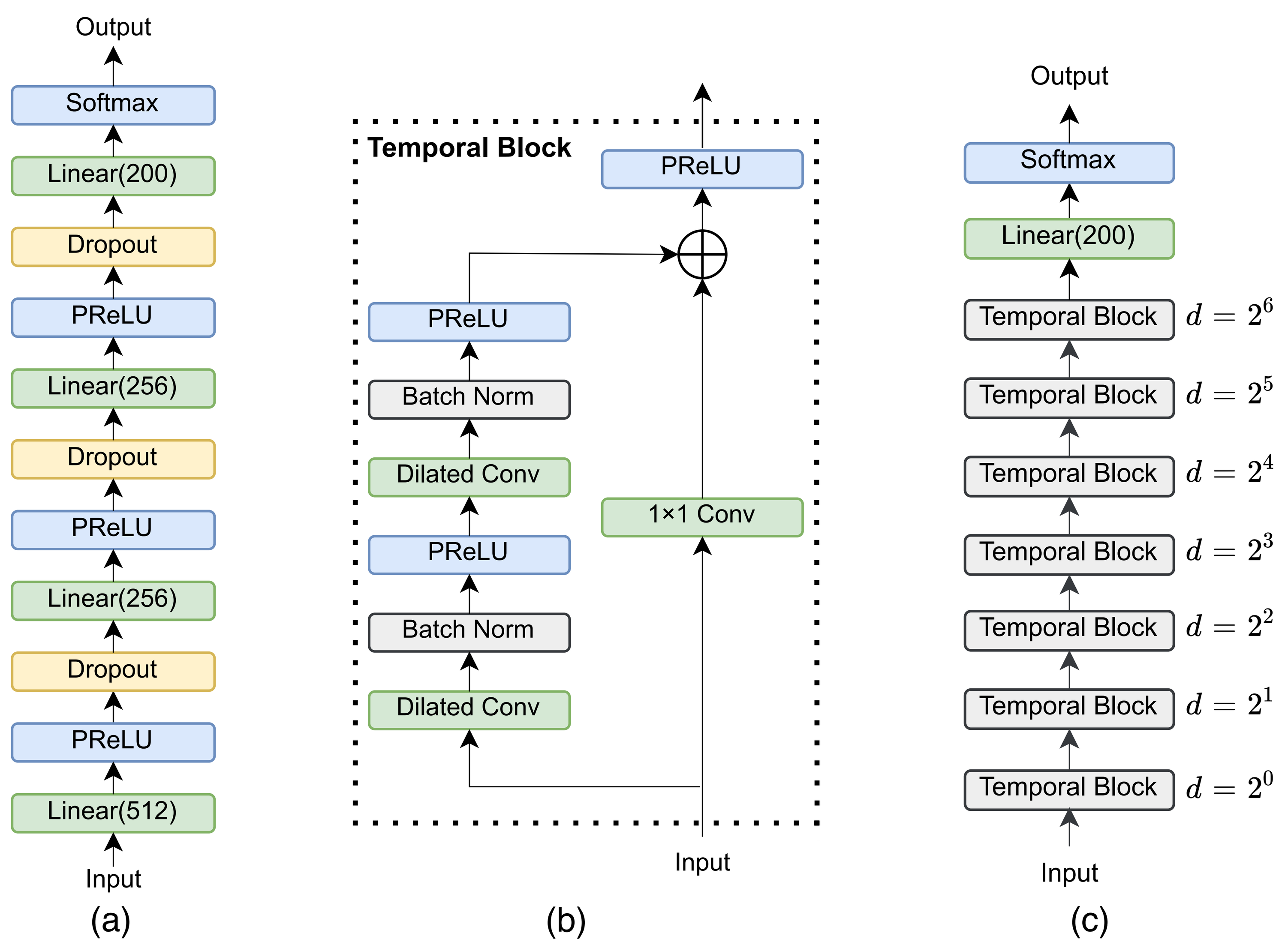

FIG. 1. The structure of our models. (a) ANN architecture for method I. (b) Temporal block in a TCN. (c) TCN architecture for method II.

Each deep learning model in method I and II was implemented and trained by Pytorch using cross-entropy loss as a loss function and Adam optimizer.

- Method III: Combination

For method III, we took the weighted average of the output predictions of methods I and II, and selected the pixel with the highest accuracy. The weight was decided to maximize the accuracy of the validation set.

RESULTS

From Table II, we can see that method III, the combination of ANN and TCN, showed the best accuracy of the three methods. Since methods I and II are substantially different and learn different features of the input, each method would have compensated deficiencies of the other and the combination of them would have increased the accuracy.

TABLE II. Accuracy of test set for each method.

	Method I	Method II	Method III
HLV	0.236	0.274	0.315
HLVK	0.351	0.401	0.458

Sky localization is measured by the 90% credible area, the smallest area enclosing 90% of the output probability, and the searched area, the smallest area that contains the actual location of the source. Of the three methods, method III showed the minimum searched area, with a medium value of 26.9 deg^2 when using three detectors (see Fig. 2). The time taken by method III to localize one GW signal was around 0.02 seconds.

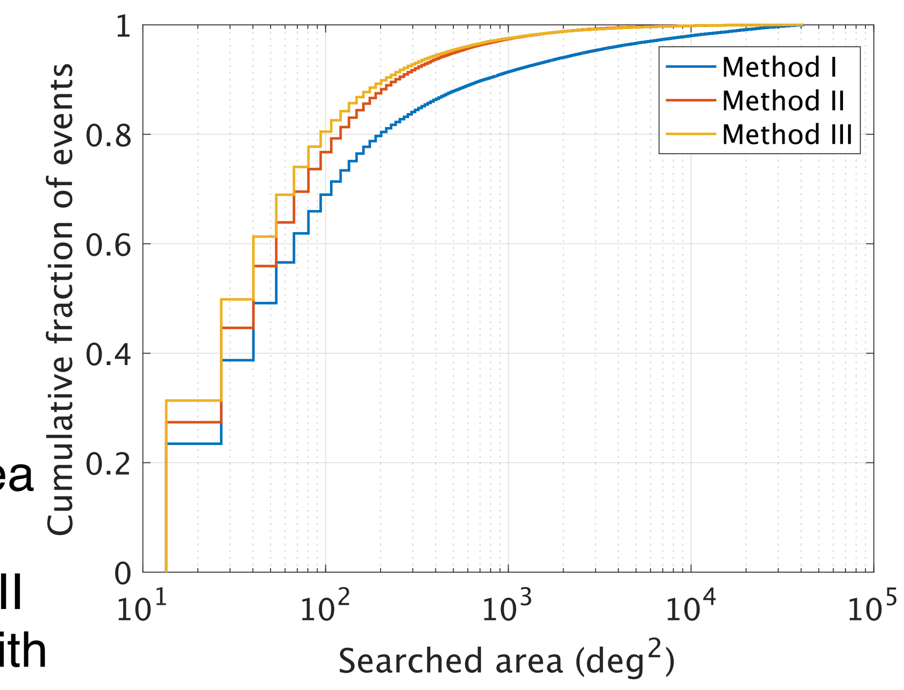

FIG. 2. Sky localization performance using three detectors (HLV) for each method.

When using four detectors, the classification accuracy was improved (Table I), and both the 90% credible area and the searched area were lowered (Fig. 3). Therefore, we confirmed that the fourth detector has a positive influence on sky localization.

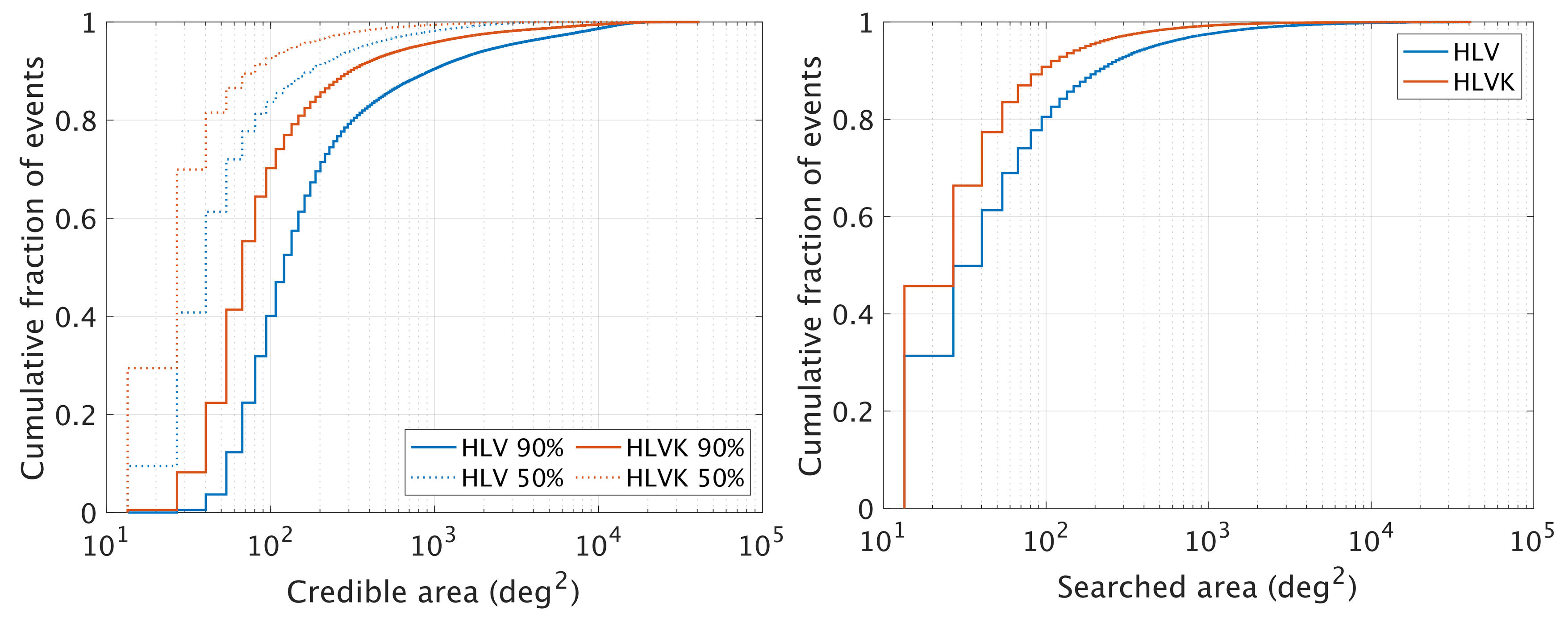

FIG. 3. Sky localization performance with and without KAGRA for method III.

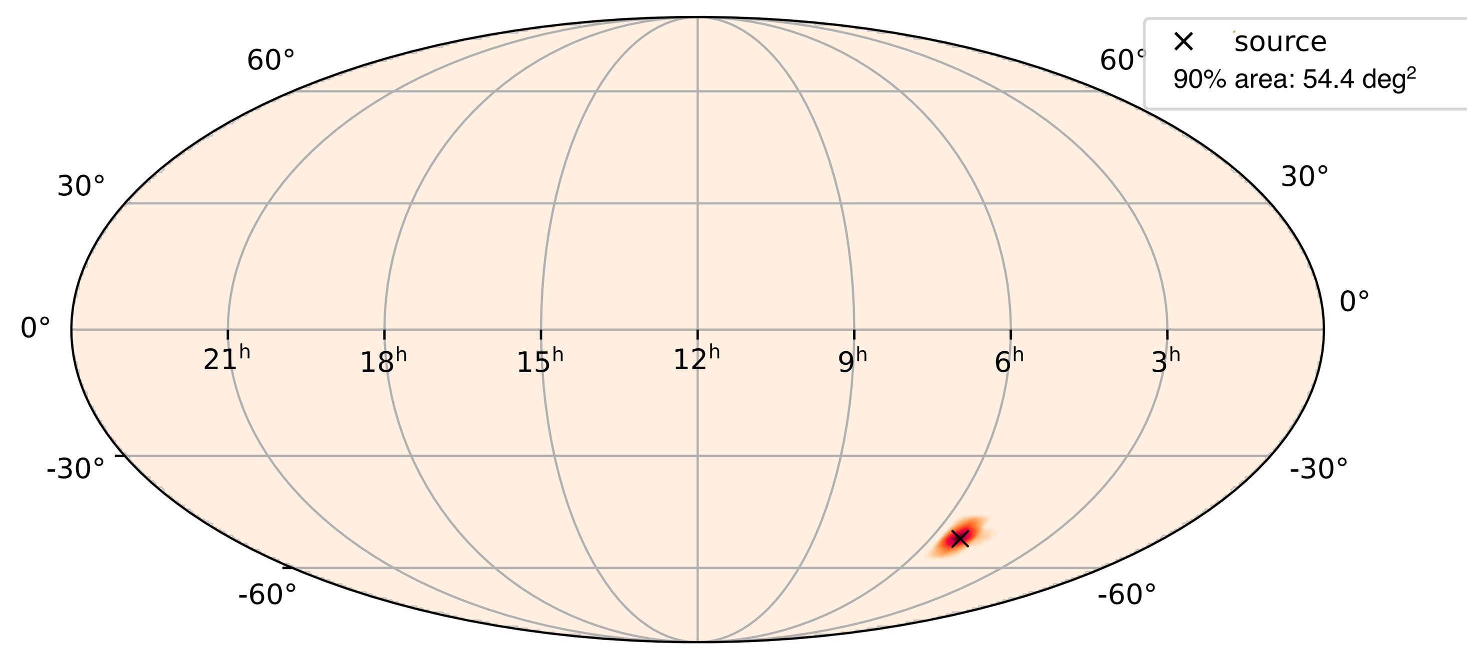

FIG. 4. Probability heatmap of localization of one test sample using four detectors.

It is a common problem in machine learning that the accuracy is low for a strain with a low SNR (Fig. 5). To solve this, we applied curriculum learning, but it did not help improve the accuracy.

To improve the accuracy for such data, in the future, we intend to apply deep learning algorithms for denoising before using our localization method. We would also like to apply this method to GWs from binary neutron star mergers and compare the results with BAYESTAR, the rapid Bayesian sky localization algorithm.

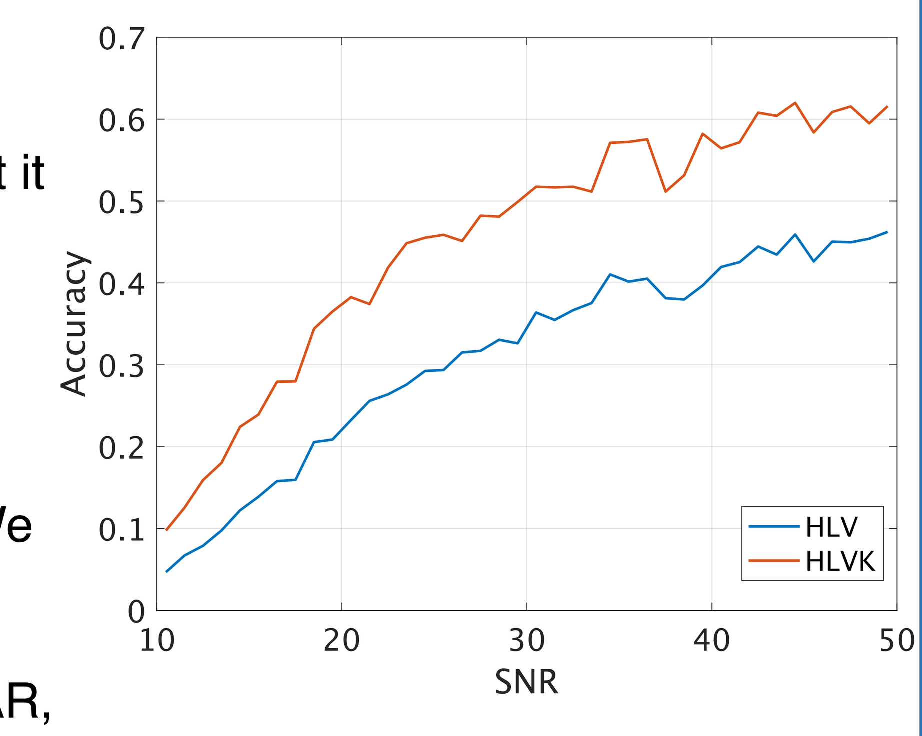

FIG. 5. Accuracy vs SNR for method III.

REFERENCES

[1] C. Chatterjee *et al.*, Phys. Rev. D **104**, 064046 (2021).