Deep Learning–based Search for Microlensing Signature from Binary Black Hole Events in GWTC-1 and -2

Kyungmin Kim1,2, Joongoo Lee3, Otto A. Hannuksela4, and Tjonnie G. F. Li4,5,6

1 Department of Physics, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Republic of Korea; kyungmin.kim@ewha.ac.kr
2 Korea Astronomy and Space Science Institute, 776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea
3 Department of Physics and Astronomy, Seoul National University, Seoul 08826, Republic of Korea
4 Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
5 Institute for Theoretical Physics, KU Leuven, Celestijnenlaan 200D, B-3001 Leuven, Belgium
6 Department of Electrical Engineering (ESAT), KU Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium

Received 2022 June 17; revised 2022 September 3; accepted 2022 September 19; published 2022 October 24

Abstract

We present the result of the first deep learning–based search for the signature of microlensing in gravitational waves. This search seeks the signature induced by lenses with masses between 10^3 and 10^5 M_☉ from spectrograms of the binary black hole events in the first and second gravitational-wave transient catalogs. We use a deep learning model trained with spectrograms of simulated noisy gravitational-wave signals to classify the events into two classes, lensed or unlensed. We introduce ensemble learning and a majority voting-based consistency test for the predictions of ensemble learners. The classification scheme of this search primarily classifies one event, GW190707_093326, into the lensed class. To verify the primary classification of this event, we also examine the median probability of the lensed class and observe that the resulting value, 0.984_{-0.342}^{+0.012}, agrees with an empirical criterion >0.6 for claiming the detection of a lensed signal. However, the uncertainty of the estimated p-value for the median probability and error, ranging from 0 to 0.1, convinces us GW190707_093326 is less likely a lensed event because it includes p ≥ 0.05 where the unlensed hypothesis is true. Therefore, we conclude our search finds no significant evidence of microlensing signature from the evaluated binary black hole events.

Unified Astronomy Thesaurus concepts: Gravitational waves (678); Astronomy data analysis (1858); Convolutional neural networks (1938); Gravitational wave astronomy (675); Gravitational microlensing (672)

1. Introduction

To date, about 100 gravitational-wave (GW) events have been identified from the data taken by the Advanced LIGO (Aasi et al. 2015) and the Advanced Virgo (Acernese et al. 2015) during the first, second, and third observing runs, referred to as O1, O2, and O3, respectively (Abbott et al. 2019, 2021a, 2021b, 2021c; Nitz et al. 2019, 2020b, 2021b, 2021a; Olsen et al. 2022). The progenitors of all events are recognized as distant compact binary mergers such as binary black holes (BBHs), binary neutron stars, and neutron star–black hole binaries; it turned out that BBH events are the majority in the population among the three source types.

Thanks to the observation of GWs, the GW astronomy era has emerged. The advent of GW astronomy has opened a new window for looking at the universe, i.e., it enables us to tackle diverse phenomena in the universe like electromagnetic-wave-based astronomy (EM astronomy) has done for centuries. For example, we could understand more about the equation of state of the interior matter of neutron stars by observing GW170817 (Abbott et al. 2018a, 2020) and could measure the Hubble parameter (Abbott et al. 2017a) by jointly observing GW170817 and GRB170817A (Abbott et al. 2017b). Besides the existing scientific investigations, many more GW avenues still await us. In this context, one such avenue with a rich EM astronomical history is the gravitational lensing of GWs (or, simply, GW lensing).

In practice, many searches have tried to look for the signature of GW lensing to date. (Broadhurst et al. 2019; Hannuksela et al. 2019; McIsaac et al. 2020; Broadhurst et al. 2020; Dai et al. 2020; Pang et al. 2020; Abbott et al. 2021e; Diego et al. 2021; Liu et al. 2021) Specifically, Hannuksela et al. (2019) and Abbott et al. (2021e) have conducted comprehensive searches exploring the signature of strong, weak, and microlensing in the 46 BBH events reported in the first and second gravitational-wave transient catalogs referred to as GWTC-1 and -2, respectively; there has been no widely accepted detection thus far.

Despite this, the forecasts of the detection rate make us expect that the detection of lensed GWs will be achievable in future observing runs: For example, Ng et al. (2018) estimated ~O(1) events per year for the strongly lensed event with the ground-based detectors’ design sensitivities reaching the redshift z ~ 1; for microlensing events, Diego et al. (2019) estimated similar rates under certain circumstances, e.g., the source is in the redshift interval 2 < z < 3, and the magnification factor is ~30.

Meanwhile, machine learning (or deep learning)–based search methods, e.g., Goyal et al. (2021) for strongly lensed events and Kim et al. (2021) for microlensed events, also have been suggested. Particularly, in Kim et al. (2021), we established a novel method utilizing deep learning for identifying the signature of microlensing in GWs (GW microlensing, hereafter). It has been discussed that GW microlensing can be recognized by its characteristic signature—beating patterns—caused by the superposition of multiple lensed GW signals arriving at the GW detector network with about O(ms) of time delays to each other (Cao et al. 2014;
Christian et al. 2018; Diego et al. 2019; Diego 2020; Jung & Shin 2019; Pagano et al. 2020; Abbott et al. 2021e; Meena et al. 2022; Seo et al. 2022). However, as discussed in the literature, the complex lens configuration embedded around macrolenses, like galaxies or galaxy clusters, or relatively weaker signature than strong lensing makes the search for GW microlensing become challenging.

Instead, in the previous work, we assumed that the strong lensing that occurs with lenses of masses between $10^5$ and $10^7 M_\odot$ might induce short time delays—comparable to that of GW microlensing—between two lensed signals: This assumption alternatively makes beating patterns appear on the waveform of BBH events as if it were GW microlensing because the desired time delay is much shorter than the typical duration of the observed BBH signals, $\lesssim 1$ s. We designed this method to seek such signatures from spectrogram images of BBH events to leverage the excellence of a state-of-the-art deep learning model, VGG-19 (Simonyan & Zisserman 2014). We supposed an arbitrary detector is optimally positioned to the source BBH’s orientation for simplicity of the proof-of-principle study. From the performance tests on two tasks, the classification of simulated signals and the regression of intrinsic and extrinsic parameters of those signals, we concluded that the method is feasible for identifying beating patterns from spectrograms of BBH events.

In this search, we revisit the BBH events examined in Hannuksela et al. (2019) and Abbott et al. (2021e) with the deep learning–based classification strategy built in Kim et al. (2021). We update the method to reflect reality a bit more, e.g., supposing the detection of such events is done via the GW detector network. Furthermore, we use the design sensitivity of the Advanced LIGO (Harry 2010) not only to mimic the irremovable noise presented in spectrograms of the BBH events but also to regard a general noise model commonly applicable for the events observed by the different detectors operated with non-identical sensitivities over the three observing runs. On top of that, we introduce ensemble learning to mitigate improperly biased predictions that might occur by using a single learner. For the initial classes on each detector’s data, predicted by the ensemble learners, we employ a majority voting-based consistency test to classify the evaluated BBH events into two primary classes, lensed or unlensed.

We figure out that one event, GW190707_093326, out of 46 is primarily classified as a lensed signal. To verify the result, we further investigate this event via the following tests: First, we examine the median probability of the lensed class of the event and observe that the result, $0.984^{+0.012}_{-0.034}$, agrees with an empirical criterion $>0.6$ for claiming the detection of a lensed signal. Second, from the uncertainty of the estimated $p$-value for the median probability and error, ranging from 0 to 0.1, we, however, are convinced that GW190707_093326 is less likely to be a lensed event because $p \geq 0.05$ where the unlensed hypothesis is true. Third, for cross-verification, we look at the Bayes factor $B_{\text{fl}}$ from Abbott et al. (2021e) and find that the $B_{\text{fl}}$ of GW190707_093326 also disfavors lensing. Therefore, we conclude the signal of GW190707_093326 is likely an unlensed signal and, consequently, we find no certain evidence of beating patterns from all evaluated BBH events, consistent with the observation made in the Bayes factor–based searches (Hannuksela et al. 2019; Abbott et al. 2021e).

We organize this paper as follows: We describe the utilization of deep learning implemented in this search in Section 2, from the configuration of training data to the application of the ensemble learning–based majority voting strategy. In Section 3, we present the search results of the event classification. Then, we provide the summary and outlook of the search in Section 4.

2. Implementation of Deep Learning

We implement one of the state-of-the-art deep learning models, VGG-19 (VGG hereafter) via PYTORCH (Paszke et al. 2019), for the identification of the GW microlensing signature. In Kim et al. (2021), we have already shown that the classification performance of the VGG model for distinguishing simulated lensed GW signals from unlensed ones is quite feasible: For example, the true-positive rates for classifying lensed signals are $>97\%$ for all considered cases.

However, it is less proper to directly use the pretrained model of the previous work because some of the considerations for preparing the training data were rather insufficient to reflect reality. For example, we supposed (i) a face-on orientation for BBH merger systems and (ii) an optimal signal-to-noise ratio $(S/N)$ for GW signals with a simpler noise model, the Detuned High Power model of the Advanced LIGO (Shoemaker 2009), in earlier work. Therefore, to mimic actual observations as much as possible, we partially change the parameter setups for the training data preparation and summarize the parameter setup used in this search in Table 1. Then we train the VGG model again with the newly prepared training data.

On the other hand, it is known in general that a single learner of any machine learning model may result in an improperly biased prediction and, for such concerns, ensemble learning can be a prescription (Géron 2017). To this end, we introduce ensemble learning for this search. Then, we build a majority voting-based classification strategy for the classes predicted by the ensemble learners.

2.1. Preparation of Training Data

We comprise the training data with spectrogram samples of mock GW signals of unlensed and lensed BBH events. First, for the generation of simulated signals, we take a similar parameter setup to that of Kim et al. (2021): For the nonprecessing unlensed GW signal in the frequency domain, $h_\text{U}(f)$, we use the IMRPhenomPv2 model (Hannam et al. 2014; Schmidt et al. 2015) with the parameters summarized in Table 1. We adopt the same prior distributions of Kim et al. (2021), i.e., the log-uniform population for the component masses and lens mass, and the uniform population for all other parameters without regard to specific prior models for the populations.\footnote{One can refer to Abbott et al. (2021f) for some practical prior models inferred from the observed GW events of interest.}

We suppose the thin-lens approximation for the lensed signals. Under the approximation, the lensed GW signal in the frequency domain, $h_\text{L}(f)$, can be described as follows:

$$h_\text{L}(f) = F(f)h_\text{U}(f),$$  

(1)

where $F(f)$ is the amplification factor in the frequency domain that determines GW lensing signatures. For the typical mass for GW microlensing, $\lesssim 10^7 M_\odot$, taking into account $F(f)$ in the wave optics is a conventional treatment (e.g., Diego et al. 2019; Seo et al. 2022). However, Takahashi & Nakamura (2003) had
shown that $|F(f)|$ in the wave optics asymptotically converges to the geometrical optics limit when a dimensionless frequency $\omega \equiv 8\pi GM_{\Delta}/c^3 \gtrsim 1$. The condition is converted to $f \gtrsim 0.1$ Hz for the considered ranges of the redshifted lens mass $M_{\Delta} = M_{\Delta}(1+z)$ in this search; we see the sensitive frequency band, $\sim$10–1000 Hz, of the ground-based detectors, corresponds to where the geometrical optic limit is valid. Hence, we take the analytic form of $F(f)$ in the geometrical optics limit provided in Takahashi & Nakamura (2003):

$$F(f) = \sqrt{\mu_+} - i\sqrt{\mu_-} e^{2\pi i f \Delta t}.$$ (2)

Here, $\mu_\pm$ are the magnification factors of two lensed signals and $\Delta t$ is the time delay of arrival times between them. To compute $\mu_\pm$ and $\Delta t$, we additionally suppose point-like lenses: For the point-mass lens model, $\mu_\pm$ and $\Delta t$ are given as

$$\mu_\pm = \frac{1}{2} \pm \frac{y^2 + 2}{2y\sqrt{y^2 + 4}},$$ (3)

$$\Delta t = \frac{4GM_{\Delta}}{c^2} \left[\frac{\sqrt{y^2 + 4}}{2} + \ln\left(\frac{\sqrt{y^2 + 4} + y}{\sqrt{y^2 + 4} - y}\right)\right],$$ (4)

respectively, where $y = (\delta D_\text{L})/(\xi_0 D_\text{S})$ is a position parameter for source which is determined by the displacement of source, $\delta$, the Einstein radius, $\xi_0 = \sqrt{(4GM_{\Delta}/c^2)D_\text{LS}D_\text{L}/D_\text{S}}$ of a point-mass lens, along with the angular distances from observer to lens, $D_\text{L}$, to source, $D_\text{S}$, and between the lens and source, $D_\text{LS}$. Now we can determine $F(f)$ by computing $\mu_\pm$ and $\Delta t$ with the parameters in Table 1; the computed values of $\mu_+, \mu_-, \Delta t$ and $\mu_\pm$ are distributed within $[1.17, 10.51], [-9.51, -0.17]$, and $[2.25$ ms, $3.52$ s], respectively.

For the three observing runs, sensitivities of the Advanced LIGO and Virgo detectors were gradually enhanced (e.g., Figure 1 of Abbott et al. 2018b) and it made the forty-six BBH events being observed in slightly different environments. Thus, we adopt the power spectral density of the Advanced LIGO’s design sensitivity (Harry 2010) not only to regard a commonly applicable noise model for the target BBH events observed in non-identical environments of detectors and observing runs but also to mimic the non-removable noise represented in the spectrogram.

We inject the simulated signals into the noise curve data acquired from the PYCBC.PSD module of the PYCBC package (Usman et al. 2016; Nitz et al. 2020a). We also constrain $S/N$ of each signal similar to Kim et al. (2021): In this search, we consider the network $S/N$, $S/N_{\text{net}}$, defined as

$$S/N_{\text{net}} = \left\{\sum_{\text{IFO}} S/N_{\text{IFO}}^2\right\}^{1/2},$$ (5)

where IFO denotes the LIGO-Hanford (H1), LIGO-Livingston (L1), and Virgo (V1) detectors and let the training data consist of spectrogram samples satisfying $10 \leq S/N_{\text{net}} \leq 50.$ By constraining $S/N_{\text{net}}$, our training samples can be prepared fairly for the evaluation on the BBH events which were identified by the detection criterion $S/N_{\text{net}} \geq 10$ at different sensitivities of the LIGO and Virgo detectors operated over the three observing runs.

We apply the constant-$Q$ transformation technique (Chatterji et al. 2004) via PYCBC.FILTER.QTRANSFORM function to the noise-added $h_{0}(f)$ and $h_{1}(f)$ signals to generate the spectrogram samples. However, we see from the observations on BBH events that the duration time of BBH signals spans less than $1$ s (for example, see Figure 10 of Abbott et al. 2019) within the sensitive frequency band of the Advanced LIGO and Advanced Virgo detectors. Therefore, we trim the time window of the spectrograms to $[-0.9$, $+0.1]$ s around the event time of each event to enhance the search accuracy. By doing so, we can save in computational expenses additionally.

We prepare 45,000 spectrogram samples for each of lensed and unlensed classes; We configure three independent subsets—training, development, and testing data—with randomly chosen 80%, 10%, and 10% of total samples, respectively, and make each subset to contain the same number of lensed and unlensed samples. We normalize the pixel values of the spectrograms to

---

**Table 1** Parameters Used for Preparing Simulated Unlensed and Lensed GW Signals

| Signal Type | Parameter | Range | Distribution |
|------------|-----------|-------|--------------|
| $h_c(f) \& h_s(f)$ | Component masses of source, $m_1$ and $m_2$ | 5–100M$_\odot$ | log-uniform |
| | Displacement of source, $b$ | 10$^{-3}$–0.5 pc | uniform |
| | R.A. of source location | 0°–360° | uniform |
| | decl. of source location | −90°–90° | uniform |
| | Polarization angle, $\psi$ | 0°–360° | uniform |
| | Inclination angle, $i$ | 0°–360° | uniform |
| | Network signal-to-noise ratio, $S/N_{\text{net}}$ | 10–50 | ... |
| $h_l(f)$ | Lens mass, $M$ | 10$^{-3}$–10$^3$M$_\odot$ | log-uniform |
| | Distance to lens, $D_L$ | 10–10$^3$ Mpc | uniform |
| | Distance from lens to source, $D_{L,S}$ | 10–10$^3$ Mpc | uniform |
| | Magnification factors, $\mu_\pm$ | $\mu_+: 1.17$–$10.51; \mu_ -: -9.51$–$-0.17$ | ... |
| | Time delay, $\Delta t$ | 2.25 ms–3.52 s | ... |

Notes. The range of network $S/N$ in the seventh row is the criterion for taking or discarding generated samples with the randomized parameters. $\mu_\pm$ and $\Delta t$ in the last three columns are the resulting values from the chosen $M$, $D_L$, and $D_{L,S}$. The prior distributions are set to be the same as those in Kim et al. (2021).

---

8 The reader can refer Section 3.1.3 of Kim et al. (2021) for the discussion about the range.

9 For the transformation, we set the frequency range as $[20$ Hz, $350$ Hz], $Q$-value range as $[16, 16]$, the step-size of time as $1/256$, and the step-size of frequency as $\log_{10}256$. Note that the values are optimal choices determined empirically in this search.
2.2. Ensemble Learning

For the implementation of ensemble learning, we repeat the training ten times with randomly chosen ten different random seeds for each training. Instead, we train all learners with the same training data by adopting the same training scheme built in the previous work (Kim et al. 2021): The batch size is set to be 128 and the maximum training epoch is set as 100. We use the Adam optimization algorithm (Kingma & Ba 2014) to optimize the training accuracy. We make the trained model returns a probability, \( r \), to the lensed class.

The loss functions taken for the error measurement \( E \) of the training is the cross-entropy function

\[
E = -r \log \hat{r} + \text{const},
\]

where \( r \) and \( \hat{r} \) are the target probability and the predicted probability of an \( i \)th training sample, respectively. The training of the ten independent learners is conducted on an NVIDIA Tesla P40 GPU. For more details, one can refer Kim et al. (2021). We present the result of performance test on the ensemble learning in Appendix A.

2.3. Majority Voting and Consistency-based Classification

To determine the initial, temporary, primary, and final classes of an event, we apply a majority voting-aided consistency test which is deployed by following hierarchical manner (see also Figure 1) with the given criteria (C):

Step 1. Determine an initial class (\( C^I \)) for the output of each ensemble learner for each detector’s data based on the probability\(^{10} \) \( r \) to the lensed class such that

- \( C^I = U \) if \( r < 0.5 \).
- \( C^I = L \) if \( r > 0.5 \).

Step 2. Decide the temporary class (\( C^T \)) of each detector’s data based on the majority voting based on the number of initial U classes \( n_U \) and the number of initial L classes \( n_L \) compared to the half of the total number of initial classes \( N_{1/2} \) such that

- \( C^T = U \) if \( n_U > N_{1/2} \).
- \( C^T = L \) if \( n_L > N_{1/2} \).
- \( C^T = R \) if \( n_U = n_L = N_{1/2} \).

Step 3. Judge the primary class (\( C^P \)) of an event based on the consistency between \( C^T \) of each detector’s data such that

- \( C^P = U \) if \( C^T = C^T \).
- \( C^P = L \) if \( C^T = C^T \).
- \( C^P = R \) if \( C^T \) and \( C^T \) are inconsistent.

Note that, for C3-2, we conservatively judge an event as U if \( C^P \)'s inconsistent with each other. Furthermore, for C4-1, we conclude \( C^P \) of an event as U if \( C^P \) is U because it is not a new discovery. On the other hand, for a detector’s data classified as R, i.e., reserved for a follow-up analysis from C2-3, we examine the mean probability \( r \) of determining \( C^P \) as either U or L via C2-3-1 or C2-3-2, respectively.

In particular, if the \( C^P \)'s of an event is judged as L from C3-3, we verify the potential new discovery by estimating the \( p \)-value based on the model established in Section 2.4 for the probability of the event. Finally, we conclude the final class of the event via C4-2 or C4-3 according to the estimated \( p \)-value.

2.4. Model for \( p \)-value Estimation

We build a \( p \)-value model from the performance test on the testing data in order to estimate the confidence of the primary classification. For the computation of \( p \)-value, we use following equation:

\[
p = 1 - \exp^{-NF},
\]

where \( N \) denotes the number of candidates satisfying the condition given in computing the false-alarm probability \( F \) defined as

\[
F = P(r^* \geq r^{|U|} | U).
\]

\( F \) means the probability of finding one or more samples having \( r^* \) greater than or equal to \( r^{|U|} \)—the probability of a target sample in the testing data—from the opposite class, i.e., the unlensed class U.

---

\(^{10}\) The probability is obtainable from the softmax activation function of the VGG model.
We present the $p$-value model built with respect to the threshold probability for the lensed samples. The shaded region around the curves shows the 1σ uncertainty on the estimated $p$-value.

We present the $p$-value model built with respect to $r'$ in Figure 2. Note that we provide the 1σ uncertainty of the $p$-value that comes from the different estimations of the ensemble learners with shaded regions. From this figure, we observe that the $p$-value becomes $<0.05$ when $r' \gtrsim 0.6$. Therefore, we set $r > 0.6$ and $p < 0.05$ as the empirical criterion for claiming the detection of a lensed signal.

3. Search Results

We present the results of the deep learning–based search for beating patterns from GW signals of the 46 BBH events in GWTC-1 and -2. We take the public strain data (32 s long and sampling rate of 4096 Hz) from the Gravitational Wave Open Science Center (Abbott et al. 2021d) to configure the evaluation data, i.e., the spectrograms of the GW signals of those events. We implement the same constant-$Q$ transformation technique and the reduced time window rule applied to the configuration of training data to the strain data, too. With this treatment, we can prepare fair spectrograms to the configuration data, i.e., the spectrograms of the GW signals of those events.

We summarize the result of the initial classification in Table 2 obtained by the criteria C1-1 and C1-2. In Figure 7 in Appendix B, we provide the initial probabilities predicted by the ensemble learners and, eventually, used for the initial classification. We tabulate the final class of each event in the last column of Table 3 together with the temporary classes obtained from the initial classes of the ensemble learners for each detector’s data and the primary class obtained by the majority voting-based consistency test for each event.

From the initial classes tabulated in Table 2, we find that five events, GW170608, GW190707_093326, GW190708_232457, GW190828_063405, and GW190930_133541, contain data that meet the criterion C2-3, i.e., $n_\ell = n_t = N_{1/2}$, so it was originally marked R in the temporary classification. But we need the temporary class to be determined as either L or U, so we test their mean probabilities and confirm that their temporary classes are all L as shown in Table 4.

We see from the confirmed temporary classes and the primary classes that 14 events meet criterion C2-2; Among them, 13 events, GW151012, GW170608, GW190412, GW190413_052954, GW190513_205428, GW190527_092055, GW190620_030421, GW190708_232457, GW190828_063405, GW190828_065509, GW190924_021846, GW190929_012149, and GW190930_133541, are C3-2 (inconsistent between the temporary classes) and the remaining 1 event, GW190707_093326, is C3-3 (all data are classified as lensed); the other 32 events correspond to criteria C2-1 and C3-1, i.e., all available detectors’ data are classified as U from the temporary classification and, eventually, the events are classified as U from the primary classifications.

For GW190707_093326, the only event classified as L from the primary classification, we compute the median probability, $\bar{r}$, from the probabilities predicted by the ensemble learners. To compute $\bar{r}$, we introduce bootstrapping which iterates the 10 probabilities 10,000 times, and we obtain $\bar{r} = 0.984^{+0.012}_{-0.034}$. Here, the error means the 90% credible interval (CI) for $\bar{r}$. We depict the $r$ of all other events with their 90% CIs in the left panel of Figure 3 together with that of GW190707_093326 for comparison. We see that there are six events, GW151012, GW170608, GW190413_052954, GW190513_205428, GW190527_092055, and GW190924_021846 partially satisfying $r \geq 0.6$ although their $r$ are computed below the empirical threshold. We observe that GW190707_093326 is the only event in which the whole range of the uncertainty agrees with the criterion. We also find from this observation that the result of the majority voting-based consistency test is consistent with the probability-based classification.

We estimate the $p$-value of each event based on the model built in Section 2.4 regarding the uncertainty in $r$. We present the resulting $p$-values in the right panel of Figure 3. From the estimated $p$-values and their uncertainties, we find that those seven events satisfying $r > 0.6$ include $p < 0.05$, even for GW190707_093326, convincing us that the unlensed hypothesis is true.

On the other hand, Abbott et al. (2021e) have also searched for beating patterns that might occur due to microlenses with masses for $M_k \lesssim 10^5 M_\odot$, which are quite similar to the consideration of this work, and the authors have defined the Bayes factor $B^\text{SIM}$ as an ultimate measure for testing the lensed hypothesis; it turned out that $\log_{10} B^\text{SIM}$ of GW190707_093326 is $-0.4$, i.e., the unlensed hypothesis is favored.

In addition to the quantitative analyses, we inspect the spectrogram of GW190707_093326. In Figure 4, we present the evaluated spectrograms11 of GW190707_093326. Note that, for the event, only two LIGO detectors, LIGO-Hanford and LIGO-Livingston, provided available data. From the spectrogram, we see that no characteristic beating pattern, i.e., neither multiple peaks nor multiple sharp nodes that might come from the possible time delay by lenses with masses for $10^3$ and $10^5 M_\odot$, is shown. In particular, even though we see from the spectrogram of LIGO-Livingston data of GW190707_093326 (the top panel of Figure 4) that the energy, represented as the brightness, of the chirp signals changes as the time evolves compared to that of LIGO-Hanford data (the bottom panel of Figure 4), it is hard to confirm the beating pattern of lensed signals like the examples shown in Kim et al. (2021).

We conclude from the follow-up analyses and a cross-verification with the Bayes factor that the signal of GW190707_093326 is likely an unlensed signal as shown in the final class of the event in Table 3. Therefore, in this search,
| Events                     | Run #1 | Run #2 | Run #3 | Run #4 | Run #5 | Run #6 | Run #7 | Run #8 | Run #9 | Run #10 |
|----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
|                            | H1     | L1     | V1     | H2     | L2     | V2     | H3     | L3     | V3     | H4     |
| GW150914                   | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW151012                   | L      | L      | U      | L      | L      | U      | U      | L      | L      | U       |
| GW151226                   | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW170104                   | L      | L      | U      | L      | L      | U      | L      | L      | U      | L       |
| GW170608                   | L      | L      | U      | L      | L      | U      | U      | U      | L      | L       |
| GW170729                   | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW170809                   | U      | L      | L      | U      | L      | U      | L      | L      | U      | L       |
| GW170814                   | U      | U      | L      | U      | U      | U      | U      | U      | U      | U       |
| GW170823                   | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190408_181802            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190412                   | L      | L      | L      | U      | L      | L      | U      | U      | L      | L       |
| GW190413_052954            | L      | U      | U      | L      | U      | L      | U      | U      | L      | U       |
| GW190413_134308            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190421_238586            | L      | L      | U      | L      | L      | L      | L      | U      | L      | U       |
| GW190424_180648            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190503_185404            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190512_180714            | U      | U      | L      | U      | L      | U      | U      | U      | U      | U       |
| GW190513_205428            | U      | L      | U      | L      | U      | U      | L      | U      | U      | U       |
| GW190514_065416            | U      | U      | L      | U      | U      | U      | L      | U      | U      | U       |
| GW190517_055101            | U      | U      | U      | U      | U      | U      | U      | U      | L      | U       |
| GW190519_153544            | U      | U      | U      | U      | U      | U      | U      | U      | L      | U       |
| GW190521                   | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190527_074359            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190602_175926            | U      | L      | U      | U      | U      | U      | U      | U      | L      | L       |
| GW190603_030421            | U      | U      | U      | U      | L      | U      | L      | U      | L      | L       |
| GW190630_185205            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190701_203306            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190706_222641            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190707_003326            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190708_232457            | U      | L      | L      | L      | L      | L      | L      | L      | L      | L       |
| GW190719_215514            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190720_000836            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190727_006333            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190728_064510            | U      | L      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190731_149036            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190803_022701            | U      | U      | U      | U      | L      | U      | U      | U      | U      | U       |
| GW190828_063405            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190828_065509            | U      | L      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190909_114149            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190910_112807            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190915_235702            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190924_021846            | U      | L      | U      | L      | U      | L      | U      | L      | L      | L       |
| GW190929_012149            | U      | U      | U      | U      | U      | U      | U      | U      | U      | U       |
| GW190930_133541            | L      | L      | L      | U      | U      | U      | U      | U      | U      | U       |

Note. The labels U and L indicate whether the event is classified as an unlensed signal or a lensed signal, respectively. Note that "Run #" in the header row denotes the index of each learner. We annotate the label of the L class by L to make them be easily distinguishable from the U class. We use this result to obtain the temporary class summarized in Table 3 or Table 4.
Table 3

| Event            | Temporary Class | Primary Class | Final Class | Event            | Temporary Class | Primary Class | Final Class | Event            | Temporary Class | Primary Class | Final Class |
|------------------|-----------------|---------------|-------------|------------------|-----------------|---------------|-------------|------------------|-----------------|---------------|-------------|
|                  | H1 L1 V1        |               |             |                  | H1 L1 V1        |               |             |                  | H1 L1 V1        |               |             |
| GW150914         | U U             |               |             | GW190503_185404  | U U             |               |             | GW190719_215514 | U U             |               |             |
| GW151012         | U L             |               |             | GW190512_180714  | U U             |               |             | GW190720_000836 | U U             |               |             |
| GW151226         | U U             |               |             | GW190513_205428  | U L             |               |             | GW190727_060333 | U U             |               |             |
| GW170104         | U U             |               |             | GW190514_065416  | U U             |               |             | GW190728_064510 | U U             |               |             |
| GW170608         | U L             |               |             | GW190517_055101  | U U             |               |             | GW190731_140936 | U U             |               |             |
| GW170729         | U U U           |               |             | GW190519_153544  | U U             |               |             | GW190803_022701 | U U             |               |             |
| GW170809         | U U U           |               |             | GW190521         | U U             |               |             | GW190828_063405 | U U             | L             |             |
| GW170814         | U U U           |               |             | GW190521_074350  | U U             |               |             | GW190828_065509 | U U             | L             |             |
| GW170818         | U U U           |               |             | GW190527_092055  | L U             |               |             | GW190909_114149 | U U             |               |             |
| GW170823         | U U             |               |             | GW190602_175927  | U U             |               |             | GW190910_112807 | ...             | U U           |             |
| GW190408_181802  | U U U           |               |             | GW190620_030421  | ...             | U L           |             | GW190915_235702 | U U             |               |             |
| GW190412         | U L U           |               |             | GW190630_185205  | ...             | U U           |             | GW190924_021846 | U L             |               |             |
| GW190413_052954  | U L U           |               |             | GW190701_203306  | U U             |               |             | GW190929_012149 | U U             |               |             |
| GW190413_134308  | U U U           |               |             | GW190706_222641  | U U             |               |             | GW190930_133541 | L'              |               |             |
| GW190421_213856  | U U             |               |             | GW190707_093326  | L'              |               |             | GW190421_180648 | ...             | U L           |             |
| GW190424_180648  | ...             |               |             | GW190708_232457  | ...             | U L           |             |                 |                 |               |             |

Note. The labels U and L indicate whether the event is classified as an *unlensed* signal or a *lensed* signal, respectively. Note that the class shown as L' denotes the confirmed temporary class, which was originally marked R according to criterion C2-3 of the classification scheme. From the list of primary classes, we see that GW190707_093326 is classified as lensed while all others are classified as unlensed from either C3-1 or C3-2. However, from the follow-up analyses, we conclude that the event is likely an unlensed one. Thus, the final class of GW190707_093326 is marked as U.
we find no certain evidence of beating patterns from all evaluated BBH events as consistent as the conclusion of Hannuksela et al. (2019) and Abbott et al. (2021e).
K.K. thanks Min-Su Shin for constructive discussion on the application of deep learning. We thank Srashiti Goyal, Thomas Dent, Young-Min Kim, John J. Oh, SangHoon Oh, and Edwin J. Son for their fruitful comments on this work. We also thank the Global Science experimental Data hub Center (GSDC) at KISTI for supporting the GPU-based computing resource. K.K. is supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT of the Korea Government (NRF-2020R1C1C1005863). O.A.H. is supported by the research program of the Netherlands Organization for Scientific Research (NWO). T.G.F.L. is partially supported by grants from the Research Grants Council of Hong Kong (Project No. 14306218), Research Committee of the Chinese University of Hong Kong, and the Croucher Foundation of Hong Kong.

This research has made use of data or software obtained from the Gravitational Wave Open Science Center (gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council, Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, and Spain. The construction and operation of KAGRA are funded by Ministry of Education, Culture, Sports, Science and Technology (MEXT), and Japan Society for the Promotion of Science (JSPS), National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea, Academia Sinica (AS) and the Ministry of Science and Technology (MoST) in Taiwan.

Appendix A
Performance Test of Ensemble Learning

Figure 5 shows the receiver operating characteristic (ROC) curves obtained by evaluating testing samples of lensed class. Figure 6 shows the results of the loss convergence test. For all ensemble learners, we see the loss for validation data set follows the convergence of loss for the training data set as desired, which shows no sign of overfitting or underfitting.

Note that the horizontal and vertical axes are the false-alarm probability $F$ given in Equation (8) and efficiency ($E$) defined as

$$E = P(r^* \geq r|L),$$

respectively. Equation (A1) means the probability of finding one or more samples having $r^*$ greater than or equal to $r^*$—the probability of a target lensed sample in the testing data—from samples of the lensed class $L$. From the ROC curves, we see that all ensemble learners can mostly correctly classify testing samples from observing $E > 0.92$ and the area $>0.99$ for all curves.

To check whether the trained model, i.e., each learner composing the ensemble learning, is either overfitted or underfitted, we examine the convergence of the loss over epochs for both training and validation data sets. As shown in Figure 6, we can see that the loss for both data sets converges as the epoch evolves, and the behavior of loss for the validation data set follows the behavior of the training data set. Therefore, we see no sign of overfitting or underfitting for all ensemble learners.
Appendix B

Probabilities from Ensemble Learners

We present the initial probabilities to the lensed class in Figure 7 for all available data of the forty-six BBH events. We use the result for the determination of the initial classification tabulated in Table 2.

![Figure 7. Probabilities of the lensed class of all data evaluated from 10 ensemble learners, referred to as “Run #.” We make the color of each cell represent the annotated value of the probability of the lensed class, such as darker blue for values close to 1 or lighter blue for values close to 0.](https://example.com/figure7.jpg)

References

Aasi, J., Abbott, B. P., Abbott, R., et al. 2015, CQGra, 32, 074011
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, ApJL, 848, L12
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2018a, PhRvL, 121, 161101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2018b, LRR, 21, 3
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019, PhRvX, 9, 031040
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2020, CQGra, 37, 054004
Abbott, B. P., Abbott, R., Adhikari, R. X., et al. 2017a, Natu, 551, 85
Abbott, B. P., Abbott, R., Abraham, S., et al. 2021a, PhRvX, 11, 021053
Abbott, B. P., Abbott, R., Abraham, S., et al. 2021d, SoftX, 13, 100658
Abbott, B. P., Abbott, R., Acernese, F., et al. 2021b, arXiv:2108.01045
Abbott, B. P., Abbott, R., Acernese, F., et al. 2021c, arXiv:2111.03606
Abbott, B. P., Abbott, R., Acernese, F., et al. 2021f, arXiv:2111.03634
Acernese, F., Agathos, M., Agatsuma, K., et al. 2015, CQGra, 32, 024001

ORCID iDs

Kyungmin Kim https://orcid.org/0000-0003-1653-3795
Joongoo Lee https://orcid.org/0000-0003-2510-4132
Otto A. Hannuksela https://orcid.org/0000-0002-5887-7137
Tjonnie G. F. Li https://orcid.org/0000-0003-4297-7365

The Astrophysical Journal, 938:157 (11pp), 2022 October 20
Kim et al.
