Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **515,** 977–983 (2022)
Advance Access publication 2022 June 21

https://doi.org/10.1093/mnras/stac1726

Deep learning unresolved lensed light curves

Mikhail Denissenya[1]★ and Eric V. Linder[1,2]

[1]*Energetic Cosmos Laboratory, Nazarbayev University, Nur-Sultan 010000, Kazakhstan*
[2]*Berkeley Center for Cosmological Physics & Berkeley Lab, University of California, Berkeley, CA 94720, USA*

Accepted 2022 June 14. Received 2022 May 9; in original form 2022 March 4

ABSTRACT

Gravitationally lensed sources may have unresolved or blended multiple images, and for time varying sources, the light curves from individual images can overlap. We use convolutional neural nets to both classify the light curves as due to unlensed, double, or quad lensed sources and fit for the time delays. Focusing on lensed supernova systems with time delays $\Delta t \gtrsim 6$ d, we achieve 100 per cent precision and recall in identifying the number of images and then estimating the time delays to $\sigma_{\Delta t} \approx 1$ d, with a $1000\times$ speedup relative to our previous Monte Carlo technique. This also succeeds for flux noise levels ~ 10 per cent. For $\Delta t \in [2, 6]$ d, we obtain 94–98 per cent accuracy, depending on image configuration. We also explore using partial light curves where observations only start near maximum light, without the rise time data, and quantify the success.

Key words: gravitational lensing: strong – methods: data analysis – methods: numerical – cosmology: observations – transients: supernovae.

1 INTRODUCTION

Gravitationally lensed Type Ia supernovae (SN Ia) should be discovered in tens and hundreds in surveys beginning in the next few years (Huber et al. 2019; Verma et al. 2019; Pierel et al. 2021; Lochner et al. 2022). These will have intriguing advantages relative to more numerous lensed quasars and other sources due to their well-defined time variation over time-scales of months, and standardizable candle nature. To increase the numbers of these useful probes of time-delay cosmology, one might attempt to use lensed systems where the images are blended or unresolved (e.g. due to lower mass galaxy lenses) and the image light curves significantly overlap.

Such overlap poses three basic observational problems: (1) recognition of the light curve as a lensed SN Ia, when the light curve does not look like a standard SN Ia, (2) determination of the number of images contributing to the light curve, and (3) estimation of the time delays between all the images. One part of the first step is identifying the light curve as being a SN Ia. This is an important initial step. Recently developed machine-learning tools using spectral information accomplish this with ∼99 per cent accuracy and can be potentially extended to classify non-supernovae sources as well (Muthukrishna, Parkinson & Tucker 2019; Davison, Parkinson & Tucker 2022); also see Boone (2021) for use of the photometric data. Our work here then picks up with identifying the SN Ia as a *lensed* SN Ia, and further carries out steps 2 and 3.

Recent articles have addressed these by considering distortions of standard supernova light curves, by expanding in 'crossing functions' (basically orthogonal polynomials) with arbitrary amplitudes (Bag et al. 2021), and by free form variation of light curves with constrained amplitudes (Denissenya et al. 2022). Both methods mentioned use Monte Carlo methods to estimate the time delays, with the second method also adding a step to identify robustly the number of images in the unresolved systems. Both demonstrated accurate measurement of the time delays for $\Delta t \gtrsim 10$ d. In this work, we turn to deep learning to accomplish this more quickly, and remove the need for separate time-delay estimation runs for each potential number of images. Other work has also investigated various aspects of unresolved light curves for lensed quasars, generally involving longer time delays than we consider, e.g. Bag et al. (2022), Biggio et al. (2021), Huber et al. (2022), Springer & Ofek (2021b, 2021a), and Shu, Belokurov & Evans (2021). We note that much of our method can be applied to more general cases of blended light curves, from a variety of transients, but we focus here on lensed SN Ia.

In Section 2, we describe the construction of the neural net, training, and test data. Section 3 presents the results for the classification of the number of images and time-delay estimations, focusing on the previous $\Delta t \gtrsim 10$ d range. We investigate higher noise systems in Section 4.1, shorter time-delay systems in Section 4.2, and the use of observational data that miss the early-time rise in Section 4.3. Discussion of results, further work, and conclusions is given in Section 5.

2 DEEP LEARNING APPROACH

The basic physical situation is of the observation of only a single blended light curve from the combination of unresolved gravitationally lensed multiple images of a time varying source. We follow the notation of Bag et al. (2021) and Denissenya et al. (2022). The observed (blended) flux in a wavelength filter j is

$$F_j(t) = \sum_{i\in\text{images}} \mu_i\, U_j(t - \Delta t_i), \tag{1}$$

where $U(t)$ is the intrinsic, unlensed source flux as a function of time (i.e. unobservable source light curve), the sum is over each image i with their individual lensing magnifications μ_i and time delays Δt_i. We focus on estimating the observable relative time delays $\Delta t_{ij} = \Delta t_i - \Delta t_j$ and determining the number of images N_{images}. Unlensed

★ E-mail: mikhail.denissenya@nu.edu.kz

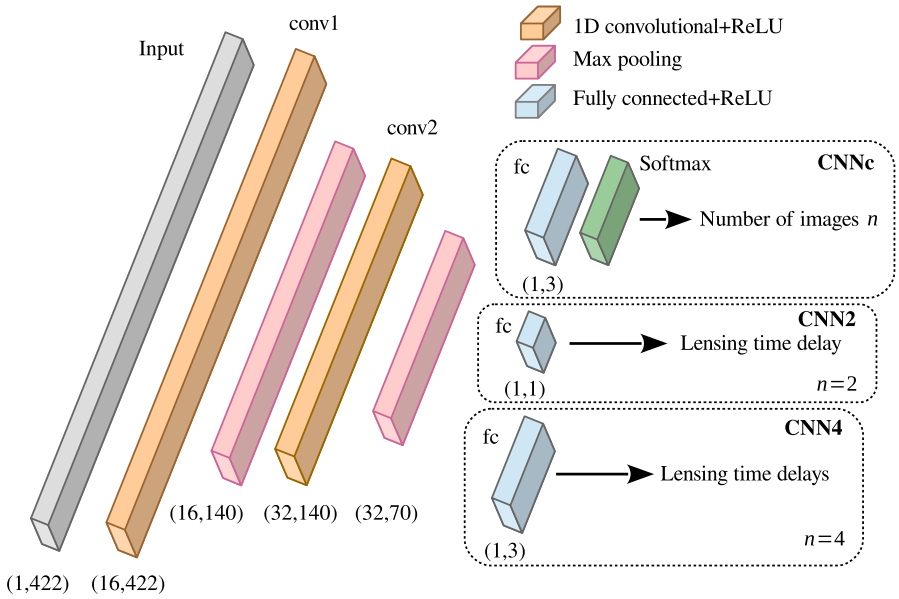

Figure 1. The unified structure of three convolutional neural networks CNNc, CNN2, and CNN4. The CNNs are designed to have the same combination of inner layers and differ in the output layers indicated enclosed in the dashed rounded boxes. The maximum probability of the softmax function of the CNNc output determines the number of images n estimated for the system. CNN2 and CNN4 neural networks estimate lensing time delays in the systems with $n = 2$ and $n = 4$ number of images, respectively. CNNc, CNN2, and CNN4 have independent training sets.

systems have one image, while multiply lensed systems have two or four images (lensing gives an odd number of images but one lensed image is generally obscured by the lens galaxy or highly demagnified).

Unlike Bag et al. (2021) and Denissenya et al. (2022), we do not input a form $U(t)$, either an expansion about a given form or a free form bounded deviation from a given form. Instead, we use a training set of generated, noisy, lensed, and unlensed SN Ia light curves and use a convolutional neural net to classify the blended light curves as arising from an $N_{\rm images}$ system. These are then fed into a convolutional neural net to perform the time-delay estimation. This same procedure could be used for any type of blended light curves, i.e. any lensed transient source, with an appropriate training set, though here we focus on SN Ia.

We explored neural networks without convolution, recurrent neural networks, and convolutional neural networks. Convolutional neural nets with two hidden convolutional layers yield the best performance; we found no significant improvement by including further hidden layers.

2.1 Training and test data

For training, we generate three data sets, each containing 10 000 systems with successive time delays between images in the range [10,14] dagnification ratios in [0.25,4], and measurement noise at the level of 5 per cent of the peak flux [so points on the rise and tail of the light curve have signal to noise much less than 20; also see Bag et al. (2021), Denissenya et al. (2022) for details], using the Hsiao supernova light-curve template (Hsiao et al. 2007) within sncosmo (Barbary et al. 2020). Our LCsimulator code used to simulate unlensed and lensed systems is publicly available in a GitHub repository.[1] Set T124 includes 1-, 2-, and 4-image systems in equal proportions, while Set T2 includes only 2-image systems and Set T4 includes only 4-image systems. We use Set T124 to train a convolutional neural network we call CNNc for classification (as unlensed, i.e. 1-image, lensed 2-image, or lensed 4-image systems). Data sets T2 and T4 train neural nets CNN2 and CNN4 respectively to accurately predict the time delays for lensed systems with the indicated number of images. We assign 80 per cent of the systems for training and employing 20 per cent for testing in each data set.

2.2 Data pre-processing

To improve computational efficiency, we take several steps to prepare the data for input. We normalize the flux measurements and the observation intervals using the corresponding ensemble (10 000 systems) maximal and minimal values. We stack flux measurements from the three (g, r, i) wavelength filters as well as the observation time instances into single input vector.

2.3 Convolutional neural network structure

In this paper, we implement three convolutional neural networks CNNc, CNN2, and CNN4 using the PyTorch high-performance deep learning library (Paszke et al. 2019). These neural networks consist of five layers described in Fig. 1 and share the same structure of inner layers consisting of alternating convolutional and max pooling layers with ReLU (rectified linear unit) activation functions. The fully connected layers form the outputs of neural networks. The output of CNNc predicts the number of images in the system. Depending on the CNNc output, we invoke either CNN2 or CNN4 to estimate lensing time delays. CNN2 has a single output number corresponding to a lensing time delay in a two-image system. If CNNc classifies the system as quadruply imaged, we employ CNN4 to estimate the corresponding relative lensing time delays, resulting in the output vector with three entries.

Each convolutional layer produces multiple copies of an input by convolving it with K kernels. In our case, the convolution turns the

[1]LCsimulator https://github.com/mdeatecl/LensedSN124imagesLCs

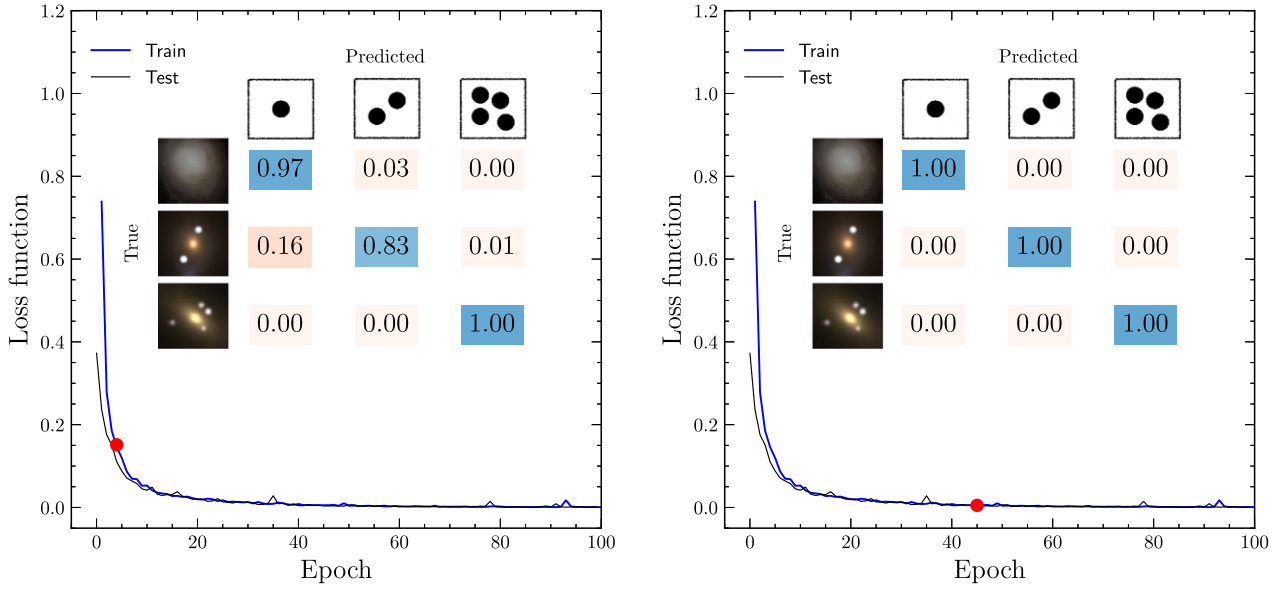

Figure 2. Total loss functions of the neural net CNNc and confusion matrices at epoch 4 (left-hand panel) and 45 (right-hand panel) obtained using the Adam algorithm (Kingma & Ba 2017) on a test set. The confusion matrix is diagonalized at epochs above 45 (matrix rows are input classes – unlensed 1-, lensed 2-, and lensed 4-image systems; columns are output classes).

input vector into a tensor with an extra depth index that comes first in a tuple defining the size of the output. The depth of the output is equal to the number of kernels K used to transform the input to output data.

We have confirmed that assigning the tasks of image number classification and estimation of lensing time delays to separate neural networks as we do is advantageous over a neural network architecture executing these task simultaneously. For example, one could use CNNc, including the CNN4 output layer dedicated to estimating the three lensing time delays in the 4-image systems. While such a neural network would still be capable of reliably identifying the number of images, the presence of unlensed and 2-image systems in the training sample but fitting for (potentially null) three time delays skews the distributions of estimated time delays even for the 4-image systems. Thus, we use the more robust architecture of CNNc followed by CNN2 or CNN4.

3 RESULTS

As described in Section 2, we train the CNNs and then validate them on data not used in training. For data from an actual survey, our goal is to productively select systems for follow-up observations with high-resolution instruments to enable the use of our time-delay estimations as a cosmological probe, covering the key range of lensing time delays 6–14 d (Goldstein, Nugent & Goobar 2019). This section assesses the purity of the image classification and time-delay estimation, and we carry out some studies of the efficiency (which will depend on survey characteristics beyond this work) in Section 4. Performance of the classification, through CNNc, can be quantified with a confusion matrix: for each input class, what fraction is classified in each output class. We find 100 per cent precision and recall in predicting the correct number of images in the 2000 lensing systems in the test set.

Fig. 2 shows the rapid convergence of CNNc towards a perfect, diagonal confusion matrix. There are no false assignments. The training process takes about ∼1 s for a single epoch (40 iterations), so about 45 s to reach the diagonal confusion matrix shown in the figure. Testing of 2000 systems takes ∼0.1 s on a 4-core CPU operating at 3.5 GHz.

Given the perfect classification, each system is unambiguously assigned to either CNN2 or CNN4 for estimation of the time delays between the two images A, B, or the four images A, B, C, and D. Fig. 3 displays those results as a histogram of the offset of the fit value from the true value, $\Delta t_{\rm fit} - \Delta t_{\rm true}$. The histograms are well-peaked and fairly symmetric.

Time-delay estimation precision $\sigma_{\Delta t} \approx 1$ d for 4-image systems and even better for 2-image systems. This is quite satisfactory: recall the time delays entering the blended light curves are in the range $\Delta t = [10, 14]$, and the observing cadence mean is 2 d. (From here on, all time delays are understood to be in units of days.) The bias, i.e. magnitude of systematic offsets from the true time delays, is $\lesssim 0.4\sigma$, so below the statistical scatter.

The CNN2 and CNN4 training process takes ∼1 s per epoch (40 iterations); the testing of 2000 systems by CNN2 takes ∼0.1 s on a 4-core CPU operating at 3.5 GHz, and testing 2000 systems by CNN4 takes roughly the same amount of time.

4 EXTENDING THE RESULTS

We next take brief looks at exploring variations, one at a time, of the fiducial data situation to extend the usefulness of this deep learning technique. Section 4.1 assesses the impact of data quality by doubling the measurement noise. In Section 4.2, we test the method by decreasing the input time delays, first down to 6 d, then all the way down to 2 d, giving even greater blending of light curves. Section 4.3 investigates the consequences of missing data from the initial rising phase of the observed light curve.

4.1 Noisier data

To investigate the impact of measurement noise on the classification and time-delay estimation, we generated another ensemble with flux noise level increased to 10 per cent, but using the same distribution of true time delays and magnifications as the fiducial data with 5 per cent

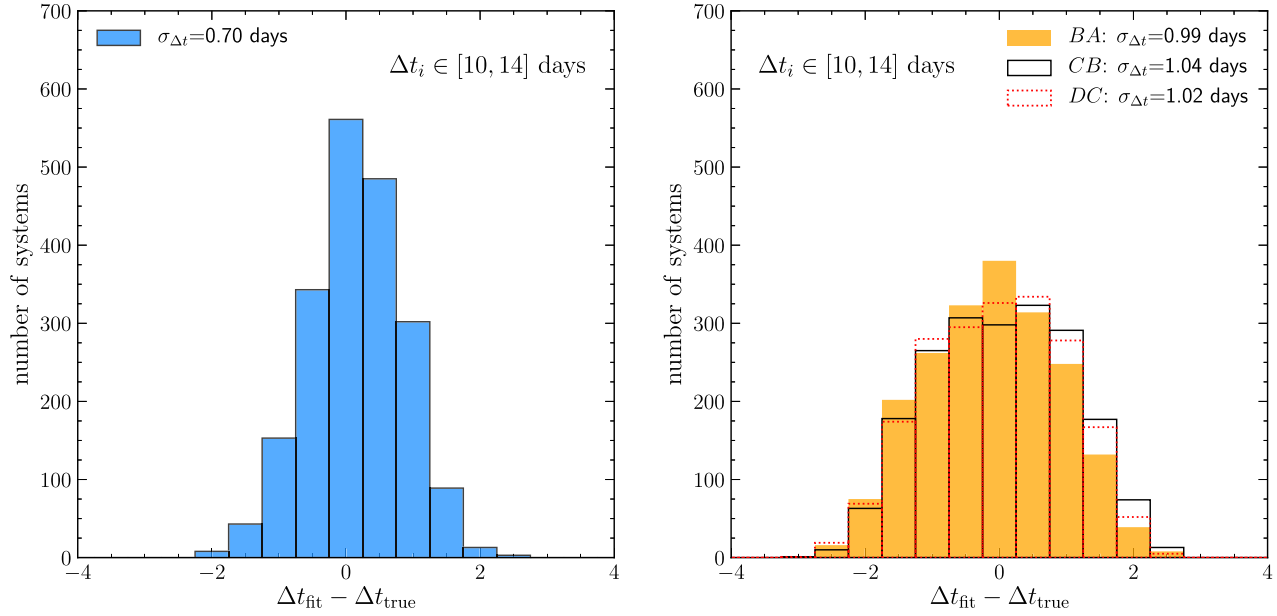

Figure 3. Histograms of time delays obtained for the test data sets having $N_{\rm images} = 2$ (left-hand panel) and $N_{\rm images} = 4$ (right-hand panel) using the trained CNN2 and CNN4 neural networks, respectively. The left-hand panel histogram shows the single time delay between the two images while the right-hand panel shows the three consecutive time delays between the four images A, B, C, and D. The histograms are quantified by the standard deviation $\sigma_{\Delta t}$, and unbiased to $\lesssim 0.4\sigma$.

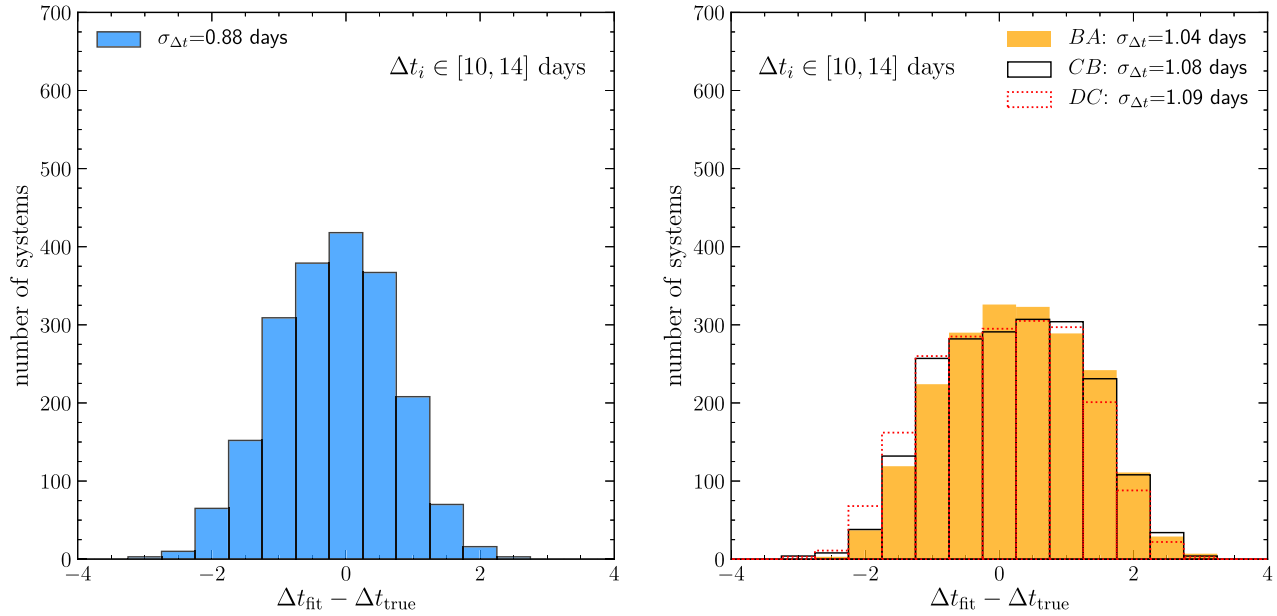

Figure 4. As Fig. 3 but for 10 per cent flux noise. Time-delay estimation remains robust.

noise level. (To allow a direct comparison of the noise impact, we have kept the random seeds the same between the 5 per cent and 10 per cent sets.) We found that a perfectly diagonal confusion matrix, i.e. 100 per cent precision and recall in predicting the number of images, still occurs in essentially the same number of iterations of CNNc, despite a noise level increased to 10 per cent. Fig. 4 shows that the distribution of predicted time delays has a slightly larger spread for the 2-image case compared to the fiducial 5 per cent noise case shown in Fig. 3, but there is little impact on the histogram width in the 4-image case. Further studies show that the classification begins to degrade at ∼20 per cent noise (much higher than expected from upcoming surveys).

4.2 Shorter time delays

In all other sections, we use a fiducial range of time delays $\Delta t = [10, 14]$ for comparison to the previous work of Bag et al. (2021) and Denissenya et al. (2022). Recall that longer time delays would

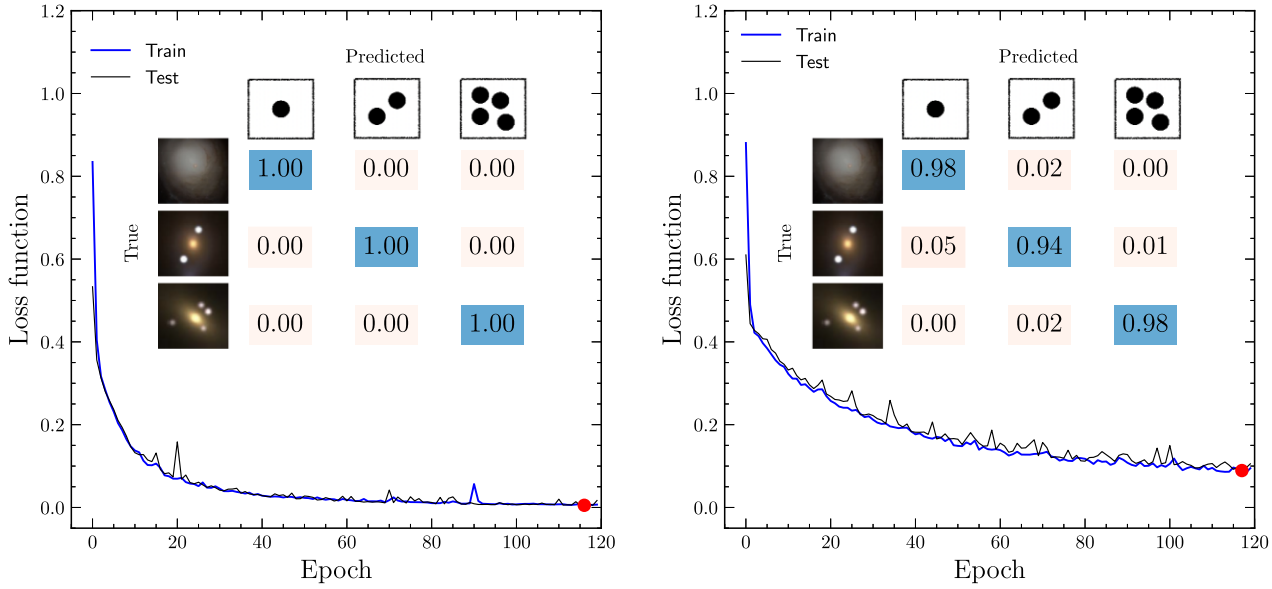

Figure 5. The loss functions of the neural CNNc net and confusion matrices evaluated for short time delays in the range [6,10] (left-hand panel) and [2,6] (right-hand panel).

give more clearly resolved light curve peaks (hence images), while the previous techniques began to falter below $\Delta t = 10$. Here, we explore how the deep-learning technique manages on shorter time delays.

For an ensemble of systems with $\Delta t = [6, 10]$ generated for training and testing with the same neural networks, the CNNc confusion matrix continues to be diagonal over this range $\Delta t = [6, 10]$. We have also verified this for the extended range $\Delta t = [6, 14]$. Finally, we pushed to even lower time delays, $\Delta t = [2, 6]$. Here, lensed 2-image systems can be mistaken for a single unlensed but broader system ~ 5 per cent of the time, while true unlensed or 4-image systems are accurately classified ~ 98 per cent of the time.

Fig. 5 shows the confusion matrices for the $\Delta t = [6, 10]$ and $\Delta t = [2, 6]$ runs. While the neural net needs to run nearly three times as long to reach the loss function minimum, the end results are, as stated above, quite good. Thus, our deep-learning technique appears robust for $\Delta t \gtrsim 6$. Fig. 6 presents the time-delay estimation histogram for the two low time-delay ranges, for the classified 2- and 4-image systems. The magnitudes $\sigma_{\Delta t}$ do not change much, though of course the fractional $\sigma_{\Delta t}/\Delta t$ increases going from the $\Delta t = [10, 14]$ case to the [6,10] and then [2,6] cases.

4.3 Missing early-time data

Observations may not capture the full lensed supernova light-curve data. We investigate the impact on the classification and time-delay estimation from missing data in the first $T_{\rm trunc}$ days after the supernova explosion. We generated multiple training sets of partial light curves by truncating the rising part of the full (untruncated) light curves. The truncation parameter $T_{\rm trunc}$ indicates the number of days missing after the ideal trigger position determined at 1 per cent of the total flux peak value for every light curve. To investigate the sensitivity of the CNNs to truncation level, we evaluated the performance for several cases: fixed $T_{\rm trunc} = 10, 20, \ldots, 80$ for each data set, and the truncation varying in the range $T_{\rm trunc} = [10, 40]$.

Fig. 7 shows the results as a function of $T_{\rm trunc}$ (we find that the case when $T_{\rm trunc}$ varies over a range simply falls in between the results corresponding to the ends of the range). The recognition of unresolved lensed images and classification of number of images remains quite robust. Defining the accuracy as the diagonal entry in the confusion matrix, we find that CNNc delivers accuracies above 0.99 for $T_{\rm trunc} \lesssim 35$, i.e. despite missing rise time data (and the time-delay estimation is insignificantly affected in this regime). Further truncation reduces the accuracy, in particular, the ability of the neural network to distinguish unlensed and 2-image partial light curves.

These results can be understood by considering that a typical supernova intrinsic rise time is $\sim$20 d, and with a single time delay of $\sim$12 d for a 2-image system, then truncation of the first $\sim$32 d means one has only data beginning near maximum flux (and extending until the supernova fades substantially, more than a month after maximum). For 4-image systems with sequential image time delays of $\sim$12 d, while one might lose the first image for extended truncations, there is still substantial flux from at least three images to later times, and such a broad light curve would look quite different from the separated maxima of a long time delay 2-image system, so the accuracy remains high.

5 CONCLUSIONS

Strongly gravitationally lensed transients carry significant cosmological information, and will increasingly be discovered by upcoming surveys. While the use of well separated, resolved, multiple images and their fluxes is established, many instances will have individual images unresolved and their light curves blended together. We build the previous work on unresolved lensed transients by using deep learning to avoid restrictive assumptions about the light-curve shape.

Taking SN Ia as an example, we show that deep learning can classify the number of images with perfect precision and recall for time delays $\Delta t \gtrsim 6$ d, accurately measure the individual time delays

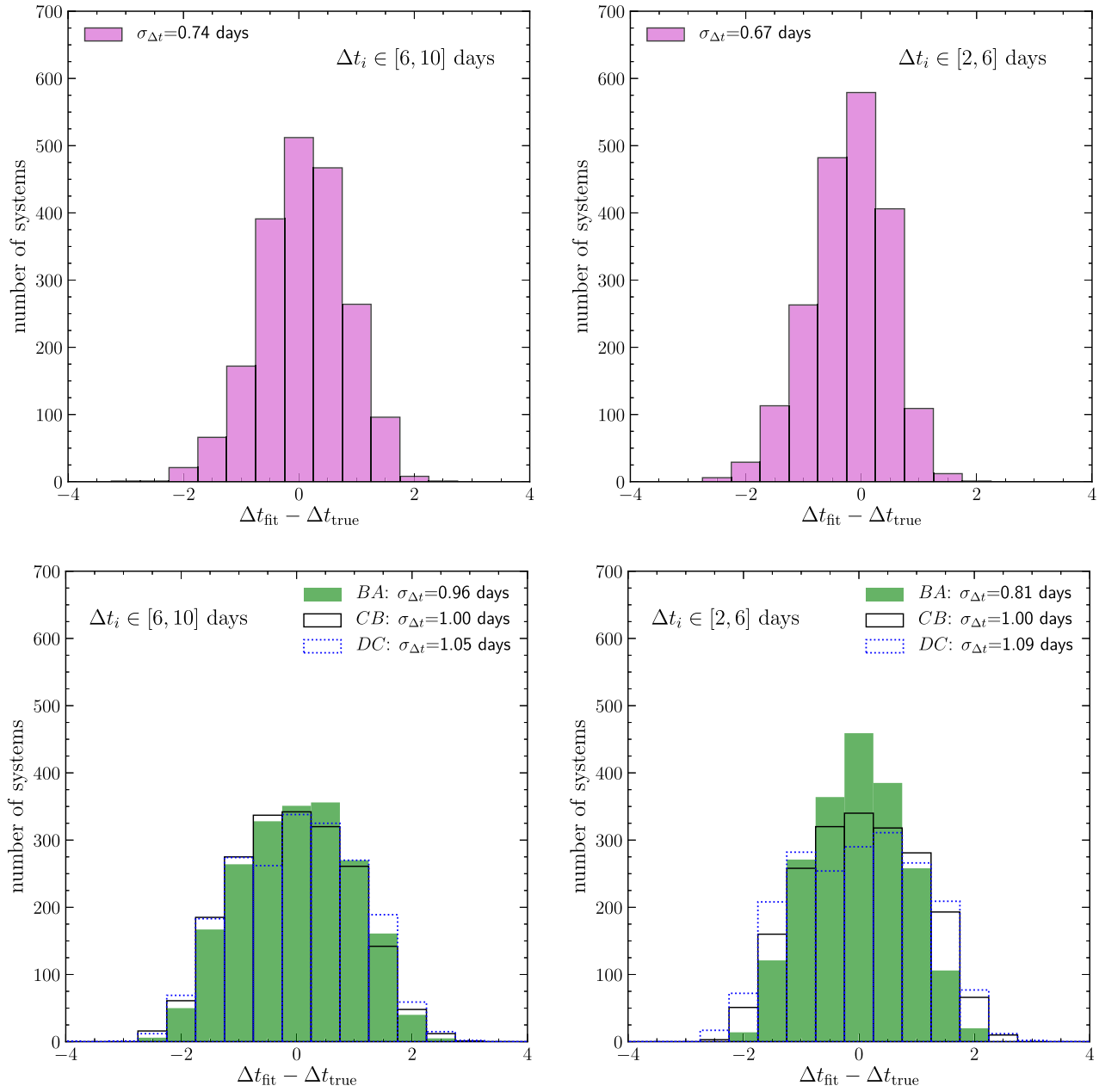

Figure 6. As Fig. 3 but for short time delays in the range [6,10] (left column) and [2,6] (right column).

(to $\sigma_{\Delta t} \approx 1$ d) without significant bias, and do so $\sim$1000$\times$ faster than previous Monte Carlo fitting techniques. A simple combination of a classification neural net CNNc and a time delay measurement neural net is found to be highly efficient and accurate.

We extended our analysis to variations of lens system and observational properties. The deep-learning implementation is robust for noise up to $\sim$ 20 per cent of maximum flux, has $\sim$ 99 per cent accuracy even when rise time data are missing up to about maximum light, and still achieves $\sim$ 95 per cent accuracy down to time delays $\Delta t \sim 2$ d.

While we have focused on SN Ia, the approach is generally applicable and can be tested for other transients in future work, including recognition of a transient as a SN Ia. A major use of identification and characterization of unresolved transients is to engage follow-up resources for cosmological probes; for example, time-delay cosmography will need not only time-delay estimation but measurement of the lens galaxy mass profile or velocity dispersion. By identifying a set of promising lensed supernovae through our deep-learning work, this can make the follow up more efficient. Further aspects to be studied include the impact of microlensing, and different intrinsic light-curve shapes, however, the speed and accuracy of this deep-learning approach should make those investigations more tractable than previous Monte Carlo fitting techniques.

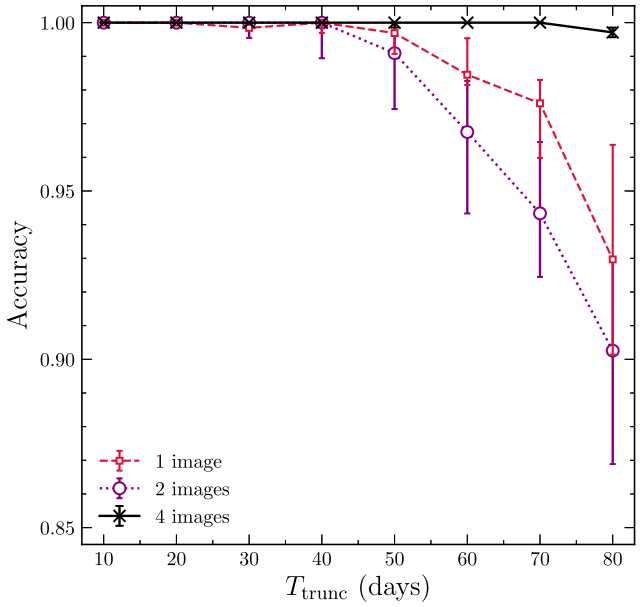

Figure 7. CNNc classification performance computed on the test sets of partial light curves at a sequence of data truncation levels. Accuracy refers to the diagonal entries in the confusion matrix.

ACKNOWLEDGEMENTS

We thank Satadru Bag, Alex Kim, and Arman Shafieloo for useful discussions. This work was supported in part by the Energetic Cosmos Laboratory. EL is supported in part by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under contract no. DE-AC02-05CH11231.

DATA AVAILABILITY

The training and test data sets for simulated light curves of 1-, 2-, and 4-image systems in Section 2 are available on the GitHub repository https://github.com/mdeatecl/LensedSN124imagesLCs4DL.

REFERENCES

Bag S., Kim A. G., Linder E. V., Shafieloo A., 2021, ApJ, 910, 65
Bag S., Shafieloo A., Liao K., Treu T., 2022, ApJ, 927, 191
Barbary K. et al., 2020, SNCosmo: Python library for supernova cosmology. Package version 2.1, Available at: https://github.com/sncosmo/sncosmo (Last accessed: 17.02.2022)
Biggio L., Domi A., Tosi S., Vernardos G., Ricci D., Paganin L., Bracco G., 2021, preprint (arXiv:2110.01012)
Boone K., 2021, AJ, 162, 275
Davison W., Parkinson D., Tucker B. E., 2022, ApJ, 925, 186
Denissenya M., Bag S., Kim A. G., Linder E. V., Shafieloo A., 2022, MNRAS, 511, 1210
Goldstein D. A., Nugent P. E., Goobar A., 2019, ApJS, 243, 6
Hsiao E. Y., Conley A. J., Howell D. A., Sullivan M., Pritchet C. J., Carlberg R. G., Nugent P. E., Phillips M. M., 2007, ApJS, 663, 1187
Huber S. et al., 2019, A&A, 631, A161
Huber S. et al., 2022, A&A, 658, A157
Kingma D. P., Ba J., 2017, preprint (arXiv:1412.6980)
Lochner M. et al., 2022, ApJS, 259, 58
Muthukrishna D., Parkinson D., Tucker B. E., 2019, ApJ, 885, 85
Paszke A. et al., 2019, in Wallach H., Larochelle H., Beygelzimer A., d'Alché-Buc F., Fox E., Garnett R., eds, Advances in Neural Information Processing Systems 32. Curran Associates, Inc.,Vancouver, British Columbia, Canada, p. 8024
Pierel J. D. R., Rodney S., Vernardos G., Oguri M., Kessler R., Anguita T., 2021, ApJ., 908, 190
Shu Y., Belokurov V., Evans N. W., 2021, MNRAS, 502, 2912
Springer O. M., Ofek E. O., 2021a, MNRAS, 506, 864
Springer O. M., Ofek E. O., 2021b, MNRAS, 508, 3166
Verma A., Collett T., Smith G. P., Strong Lensing Science Collaboration, the DESC Strong Lensing Science Working Group, 2019, preprint (arXiv:1902.05141)

This paper has been typeset from a TEX/LATEX file prepared by the author.