Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **514,** 5159–5177 (2022) https://doi.org/10.1093/mnras/stac1691
Advance Access publication 2022 June 20

The dark energy survey 5-yr photometrically identified type Ia supernovae

A. Möller [1,2★] M. Smith,[3] M. Sako,[4] M. Sullivan [5] M. Vincenzi [6] P. Wiseman [5] P. Armstrong,[7] J. Asorey [8] D. Brout [9] D. Carollo,[10] T. M. Davis [11] C. Frohmaier [5,12] L. Galbany [13,14] K. Glazebrook,[1] L. Kelsey [5,12] R. Kessler,[15,16] G. F. Lewis [17] C. Lidman [7,18] U. Malik,[7] R. C. Nichol,[19] D. Scolnic [6] B. E. Tucker,[7] T. M. C. Abbott,[20] M. Aguena [21] S. Allam,[22] J. Annis,[22] E. Bertin [23,24] S. Bocquet,[25] D. Brooks,[26] D. L. Burke,[27,28] A. Carnero Rosell [21,29,30] M. Carrasco Kind [31,32] J. Carretero [33] F. J. Castander,[34,35] C. Conselice,[36,37] M. Costanzi,[38,39,40] M. Crocce [34,35] L. N. da Costa,[21,41] J. De Vicente [8] S. Desai,[42] H. T. Diehl,[22] P. Doel,[26] S. Everett,[43] I. Ferrero [44] D. A. Finley,[22] B. Flaugher,[22] D. Friedel,[31] J. Frieman,[16,22] J. García-Bellido [45] D. W. Gerdes,[46,47] D. Gruen,[25,48] R. A. Gruendl,[31,32] J. Gschwend,[21,41] G. Gutierrez,[22] K. Herner,[22] S. R. Hinton [11] D. L. Hollowood [43] K. Honscheid,[49,50] D. J. James,[9] K. Kuehn,[51,52] N. Kuropatkin,[22] O. Lahav,[26] M. March [4] J. L. Marshall,[53] F. Menanteau,[31,32] R. Miquel,[33,54] R. Morgan,[55] A. Palmese [56] F. Paz-Chinchón,[31,57] A. Pieres [21,41] A. A. Plazas Malagón [58] A. K. Romer,[59] A. Roodman,[27,28] E. Sanchez,[8] V. Scarpine,[22] M. Schubnell,[47] S. Serrano,[34,35] I. Sevilla-Noarbe,[8] E. Suchyta [60] G. Tarle,[47] D. Thomas [12] C. To [49] and T. N. Varga[61,62]

Affiliations are listed at the end of the paper

Accepted 2022 June 14. Received 2022 June 13; in original form 2022 January 25

ABSTRACT

As part of the cosmology analysis using Type Ia Supernovae (SN Ia) in the Dark Energy Survey (DES), we present photometrically identified SN Ia samples using multiband light curves and host galaxy redshifts. For this analysis, we use the photometric classification framework SuperNNovatrained on realistic DES-like simulations. For reliable classification, we process the DES SN programme (DES-SN) data and introduce improvements to the classifier architecture, obtaining classification accuracies of more than 98 per cent on simulations. This is the first SN classification to make use of ensemble methods, resulting in more robust samples. Using photometry, host galaxy redshifts, and a classification probability requirement, we identify 1863 SNe Ia from which we select 1484 cosmology-grade SNe Ia spanning the redshift range of $0.07 < z < 1.14$. We find good agreement between the light-curve properties of the photometrically selected sample and simulations. Additionally, we create similar SN Ia samples using two types of Bayesian Neural Network classifiers that provide uncertainties on the classification probabilities. We test the feasibility of using these uncertainties as indicators for out-of-distribution candidates and model confidence. Finally, we discuss the implications of photometric samples and classification methods for future surveys such as Vera C. Rubin Observatory Legacy Survey of Space and Time.

Key words: methods: data analysis – surveys – supernovae: general – cosmology: observations.

1 INTRODUCTION

To fully exploit the power of current and future time-domain surveys, it is necessary to classify astrophysical objects using only photometry. Surveys such as the Supernova Legacy Survey (SNLS), *Sloan Digital Sky Survey* (*SDSS*) SN Survey (SDSS-II), Pan-STARRS (PS1), and the Dark Energy Survey (DES) have discovered thousands of supernovae (SNe) but the majority have not been spectroscopically classified (Astier et al. 2006; Frieman et al. 2008; Bernstein et al. 2012; Rest et al. 2014; Foley et al. 2018; Sako et al. 2018; Smith et al. 2020). Photometric classification will be particularly crucial for the upcoming Legacy Survey of Space and Time (LSST) at the Vera C. Rubin Observatory, which is expected to discover up to 10^7 SNe over the next decade (LSST Science Collaboration 2009).

The Dark Energy Survey Supernova programme (DES-SN) obtained photometry of more than 30 000 candidate SNe over its five years of operation. These include thousands of high-redshift SNe Ia, of which only several hundred have been spectroscopically classified. The first three years of the DES-SN detected and spectroscopically classified 251 SNe Ia (Smith et al. 2020). Together with low-redshift SNe from the Harvard–Smithsonian Center for Astrophysics surveys (CfA3, CfA4; Hicken et al. 2009, 2012) and the Carnegie Supernova Project (CSP; Contreras et al. 2010; Stritzinger et al. 2011), these SNe

★ E-mail: amoller@swin.edu.au

were used to constrain cosmological parameters (Dark Energy Survey 2019). The DES-SN candidate sample also contains other types of transients that have been used for astrophysical and cosmological studies: core-collapse SNe (de Jaeger et al. 2020), superluminous SNe (SLSNe; Smith et al. 2018; Angus et al. 2019; Inserra et al. 2021), rapidly evolving transients (Pursiainen et al. 2018; Wiseman et al. 2020b) and 'peculiar' events (Gutiérrez et al. 2020; Grayling et al. 2021).

To classify SNe without spectroscopy, a number of methods have been developed to classify them using their light curves, i.e. their observed brightness evolution in different filters. Due to their cosmological use, much work has focused on disentangling SNe Ia from other SN types. The majority have been trained and tested on simulations, with only a handful applied to large SN surveys (Sako et al. 2011; Möller et al. 2016; Möller & de Boissière 2019; Muthukrishna et al. 2019; Villar et al. 2019, 2020). Several photometric classifiers have been developed and incorporated into the SNIa-cosmology analysis pipeline PIPPIN (Hinton & Brout 2020), including SNIRF (based on the architecture developed by Dai et al. 2018), SUPERNNOVA (Möller & de Boissière 2019), and SCONE (Qu et al. 2021).

In this work, we use the non-parametric framework SUPERNNOVA (SNN; Möller & de Boissière 2019) to obtain photometrically classified SN Ia samples from DES-SN. SNN has several strengths it: (i) requires only photometric information (fluxes and time) for classification, (ii) does not rely on the extraction of features, (iii) can be trained to classify any type of transient event, (iv) can use redshifts to improve accuracy, (v) has been thoroughly tested using simulations, (vi) includes algorithms that assign uncertainties to classification probabilities such as Bayesian Neural Networks (BNNs), and (vii) is already being applied to real survey data, including early light-curve classification in alert streams (FINK broker; Möller et al. 2021).

Photometrically classified SN Ia samples have started to be used in cosmology. First constraints on the cosmic expansion using data from SDSS-II and PS1 have shown the feasibility of using these samples for cosmology and their competitive constraining power on the Dark Energy (Sako et al. 2011; Hlozek et al. 2012; Campbell et al. 2013; Jones et al. 2017, 2018). Most of these results use the Bayesian Estimation Applied to Multiple Species method (BEAMS; Kunz, Bassett & Hlozek 2007) and its extension 'BEAMS with Bias Corrections' (BBC; Kessler & Scolnic 2017). These methods incorporate classification probabilities of SNe Ia into the analysis, thus requiring accurate classification probabilities. Recent work estimates the contamination for cosmological constraints in the DES-SN sample using SNNat less than 1.4 per cent (Vincenzi et al. 2022). Aside from cosmology, photometrically classified samples with SNN have also been used to study SN Ia rates (Wiseman et al. 2021).

This paper is organized as follows: We introduce the DES survey and DES-SN candidate sample in Section 2. In Section 3, we present pre-processing needed for accurate classification, SUPERNNOVA, realistic simulations, training and classification mechanisms and their metrics. In Section 4, we select photometrically classified SNe Ia using host galaxy redshift information together with multiband photometry. We explore the use of BNNs for classification in Section 5. Finally, in Section 6, we discuss our results and their implications for future surveys such as LSST.

2 DES-SN 5-YR

The Dark Energy Survey (DES) was a 6-yr photometric survey that used the Dark Energy Camera (DECam; Flaugher et al. 2015) on the Victor M. Blanco telescope in Chile to survey 5000 deg^2 of the Southern hemisphere. For time-domain science, DES imaged ten 3-deg^2 in the *griz* filters during the first 5 yr (Abbott et al. 2018). Eight of these ten fields (X1, X2, E1, E2, C1, C2, S1, and S2) were observed to a single-visit depth of $m \approx 23.5$ mag ('shallow fields'), and the other two 'deep fields' (X3,C3) were observed to a depth of $m \approx 24.5$ mag.

2.1 DES-SN candidate sample

Transients were identified using the DES Difference Imaging Pipeline DIFFIMG (Kessler et al. 2015) coupled with a machine-learning algorithm (Goldstein et al. 2015) to reduce artefacts. A candidate SN is defined from the difference image measurements by requiring at least two detections with a signal-to-noise ratio (SNR) larger than five in any filter. This criteria is designed to remove artefacts and asteroids.

Each DES-SN candidate was originally associated with a host galaxy using the shallower SVA survey, created from DES Science Verification data. For the DES-SN analysis, we use deep co-adds in Wiseman et al. (2020a). The major source of host galaxy redshift information was the Australian Dark Energy Survey (OzDES) programme obtaining spectra with the 2dF fibre positioner and AAOmega spectrograph on the 3.9-m Anglo-Australian Telescope (Yuan et al. 2015; Childress et al. 2017; Lidman et al. 2020). SN hosts in OzDES were observed up to a limiting r magnitude of ≈24. Further details on host galaxy association can be found in Gupta et al. (2016), Vincenzi et al. (2020).

For the 31 636 candidates, 29 113 have an identified host and 11 350 have a spectroscopic redshift (∼30 per cent of the candidates).

A sub sample of candidates were selected for real-time spectroscopic follow-up observations for classification. For the first 3 yr of the survey, the spectroscopically classified sample is presented in Smith et al. (2020). In this work, we use for comparison a preliminary spectroscopic sample containing additional classifications from the full 5 yr of DES-SN. This sample contains 415 spectroscopically confirmed SNe Ia (including all 251 spectroscopically classified SNe Ia from the DES-SN 3-yr analysis), 84 core-collapse SNe, 2 peculiar SNe Ia, 20 SLSNe, 55 AGN, 1 Tidal disruption event (TDE), and 2 M-stars. We highlight that this spectroscopically classified sample is not complete (Kessler et al. 2019b) and does not represents the true abundances of different transients in nature.

In this work we use the fluxes and uncertainties obtained from DIFFIMG (Kessler et al. 2015) for the DES-SN candidate sample.

2.2 Filtering multiseason and other transients

The DES-SN 5-yr candidate sample contains not only supernovae but also astrophysical events such as fast transients and AGNs. These events, called out-of-distribution (OOD) or anomalies, can be hard to characterize and thus simulate, therefore photometric classifiers are usually not trained to identify them.

To reject fast, very low SNR transients or transients that have a limited photometric sampling (e.g. transients occurring near the end or beginning of the observing season), we select only transients that have at least 3 nights with a detection that has passed the DES Real/Bogus image classifier (Goldstein et al. 2015).

To reduce the number of slowly evolving transients that span several observing seasons or multiseason candidates (e.g. AGNs) and spurious detections we make use of two selection criteria. First, we compute the ratio between number of epochs with detections that pass the Real/Bogus classifier, and the total number of epochs with

detections. To reject light curves with long variability periods, we require this ratio to be the same as in the real-time classification pipeline in Smith et al. (2020). Secondly, we remove artefacts and transients that have detections in multiple observing seasons. We note that this cut can remove real supernovae, i.e. multiple SNe very close-by in the same galaxy, and is not 100 per cent efficient.

With this filtering, the sample is reduced from 31 636 to 14 070 candidates. This reduces the number of candidates and their contamination; however, some residual AGN and other types of SNe remain. We find that this sample includes 405 spectroscopically classified SNe Ia (247 of which are in the DES-SN 3yr sample), 83 core-collapse SNe, 2 peculiar SNe Ia, 19 SLSNe, 37 AGN, 1 TDE, and 1 M-star.

2.3 Selection requirements (cuts)

We apply a series of selection cuts on both the quality of the light curves and the quality of the redshifts. A thorough review on these cuts and their impact on systematics can be found in Vincenzi et al. (2022).

2.3.1 Loose selection cuts

We select transients that have redshifts obtained from spectra from either the SN or its host galaxy (Lidman et al. 2020) using quality tags in Vincenzi et al. (2020). In this work, we also include lower resolution redshifts from PRIMUS since they are precise enough for photometric classification.[1] After this selection cut we obtain 6635 SN candidates.

Furthermore, we restrict these redshifts to be within the range of the SNe Ia expected for DES-SN and thus in our simulations, $z \in [0.05,1.3]$. This cut also removes stars in our catalogues.

We fit the light curves using the SALT2 model (Guy et al. 2007). We require that: (i) at least two filters have at least one observation with SNR larger than 5, (ii) at least one photometric measurement before peak brightness t_0, and (iii) at least one photometric point ten days after peak brightness.

We select a sample of 2381 light curves that satisfy these sampling criteria and have a SALT2 fit that converges and is within SALT2 model boundaries for stretch, $x1 \in [-4.9, 4.9]$ and colour $c \in [-0.49, 0.49]$. We photometrically classify these candidates in the following. This sample contains a subsample of spectroscopically classified candidates which we will use as a reference: SNe Ia: 366 (DES-SN 3 yr: 228), CC 13, SLSN 2, AGN 3. The SALT2 parameters (amplitude, stretch, and colour) are not used by SNN.

2.3.2 JLA-like cuts

We will consider an additional set of cuts after photometric classification based on the criteria in Vincenzi et al. (2022). They will only be applied when specified.

These cuts are designed to select cosmology-grade SNe Ia and are based on those from the Joint Light-curve Analysis: $-3.0 < x_1 < 3.0$, $-0.3 < c < 0.3$, and $\sigma_{x_1} < 1$ and $\sigma_{t0} < 2$ (Betoule et al. 2014). Where $c, x_1, \sigma_{t0}, \sigma_{x_1}$ are estimated using SALT2 and represent colour, stretch, and uncertainty on t_0 and x_1, respectively. These cuts are implemented in SN Ia cosmology analyses to restrict SNIa parameters to the valid model range, and to reject peculiar SNIa. We also use a SALT2 fit probability >0.001 selection.

[1]The redshifts from the PRIsm MUlti-object Survey (PRIMUS) were obtained using the Inamori Magellan Areal Camera and Spectrograph camera on the Magellan I Baade 6.5 m telescope (Coil et al. 2011). They are less accurate and they have a higher rate of catastrophic failure, thus not suitable for cosmological constraints.

3 PHOTOMETRIC CLASSIFICATION

We use the photometric classification algorithm SUPERNNOVA (SNN) to select SN Ia from the DES-SN 5-yr candidate sample that pass loose selection cuts. We introduce pre-processing necessary for accurate photometric classification of our DES-SN 5-yr data (Section 3.1). We generate realistic simulations of the DES-SN survey to train and test our photometric classification method (Section 3.2) and the framework SNN (Section 3.3). We evaluate performance and find the best configuration for our framework using small simulations (Section 3.4). We then train optimized models for photometric classification of the DES-SN 5-yr sample using larger simulations (Section 3.5).

3.1 DES-SN data pre-processing

For accurate photometric classification, the simulations used to train the models and the data to be classified should be similar. While light-curve simulations strive to resemble survey data, pre-processing of the survey data is required to assure this.

First, DES-SN data were taken over five consecutive seasons. Each DES season represented about five months of observations per year. SNe last only for months, thus are only detected in a subset of this photometry. In our simulations (see Section 3.2), supernovae are simulated within a rest-frame time-span, e.g. −30 d before to 100 d after peak luminosity. To select an equivalent time window in the DES-SN 5-yr data, we first obtain an estimated time of peak brightness (t_0) using the SuperNova ANAlysis software (SNANA; Kessler et al. 2009). This t_0 estimate is not obtained using SALT2 (Guy et al. 2007), but instead based on max flux in region of dense detections to avoid pathological estimates from a single pathological flux in another season. Once the peak has been determined for each light curve, we select and classify photometric points within an observed time-window around the light-curve peak of [−30, 100] d.

Light curves may contain photometry that has been flagged as flawed. We require that SNNdiscard photometry that is not reliable using the bitmap flag provided by SOURCE EXTRACTOR (Bertin & Arnouts 1996) and DIFFIMG (Kessler et al. 2015). These photometric outliers are not present in the simulations used to train our photometric classifier. This is in particular important when using normalization schemes, which will be introduced in Section 3.3.1, since they use maximum fluxes to normalize the light curves. If that maximum flux comes from a bad photometric point, the light curve will be distorted and therefore classification will not be accurate. This photometry quality criteria reduces the number of photometric measurements by 6 per cent but keeps the number of transients unchanged.

3.2 Simulations of the DES-SN survey

SNNis used with simulations from the supernova analysis software (snana Kessler et al. 2009) and within the PIPPIN orchestration framework (Hinton & Brout 2020). The simulations incorporate information from DES-SN observations (PSF, sky noise, zero-point), with detection efficiencies versus SNR estimated on fake SNe that were overlaid on images and processed with DIFFIMG. Simulations include SNe that have partial light curves due to season boundaries or observing gaps imitating realistic weather conditions. Detailed

Table 1. Simulations used for training and testing SNN. Columns indicate simulation name, approximate number of light curves generated and number of light curves when balancing simulations to have the same number of normal Type Ia and other SNe.

Simulation name	Number of light curves (10^6)	Balanced number of light curves (10^6)
TRAIN-SIM	4.5	3.63
S-TRAIN-SIM	2.0	1.4
TEST-SIM	0.8	Not applicable

information on the inputs necessary to obtain realistic DES-SN simulations can be found in Kessler et al. (2019b). We also make use of recent updates in the library of simulated host galaxies for DES-SN as introduced by Vincenzi et al. (2020). This host galaxy library includes the dependence of SN rates on galaxy properties such as stellar mass and galaxy star formation rate.

We simulate a variety of SNe using volumetric rates and input parameters as described in Vincenzi et al. (2020). Our simulations are performed over a redshift range $0.05 < z < 1.3$. These simulations contain normal SNe Ia, peculiar SNe Ia, and core-collapse SNe.

Normal SNe Ia are generated using the SALT2 SED model presented in Guy et al. (2007), trained for the JLA sample (Betoule et al. 2014) and extended to UV and IR wavelengths (Pierel et al. 2018) to improve the redshift coverage of our simulated SNe. Volumetric rates from Frohmaier et al. (2019) are used. The intrinsic stretch and colour distributions are taken from Scolnic & Kessler (2016) and we use the G10 intrinsic scatter model from Kessler et al. (2013) based on Guy et al. (2010). Peculiar SNe Ia include SN91bg-like (Kessler et al. 2019a) and SNe Iax (Jha 2017) with models updates in Vincenzi et al. (2022).

We make use of three different core-collapse SN template collections: V19 (Vincenzi et al. 2019), J17 (Jones et al. 2017), and templates used in the Supernova Photometric Classification Challenge (SPCC; Kessler et al. 2010). The main differences between these templates include: the number of SNe used to create them, the rates used, and the interpolation methods and wavelength coverage. Detailed information on these templates can be found in Vincenzi et al. (2019).

Our baseline simulations, and used unless specified, are generated using V19 core-collapse SN templates. Relative core-collapse SN rates are given by Li et al. (2011) updated in Shivvers et al. (2017) and the total rate is assumed to follow the cosmic star formation history presented in Madau, Weisz & Conroy (2014) normalized by the local SN rate of Frohmaier et al. (2019).

We generate different simulations to train (TRAIN-SIM and a smaller S-TRAIN-SIM for computing efficiency of certain evaluation tasks) and test (TEST-SIM) SNNas shown in Table 1. For training, after generating the simulation, we randomly trim the simulation to ensure a balanced training sample, with the same number of normal SNe Ia and non-normal Ia (core-collapse SNe and peculiar SNe Ia). Volumetric rates guarantee that the mixture of non-Ia SNe is consistent with measured rates. We note that the size of the S-TRAIN-SIM training set is the same as the complete sample used in Möller & de Boissière (2019). Having defined our simulated samples we now turn to methods of classifying them.

3.3 SuperNNova (SNN)

SuperNNova (Möller & de Boissière 2019) is a deep learning framework for light-curve classification. It makes use of fluxes and

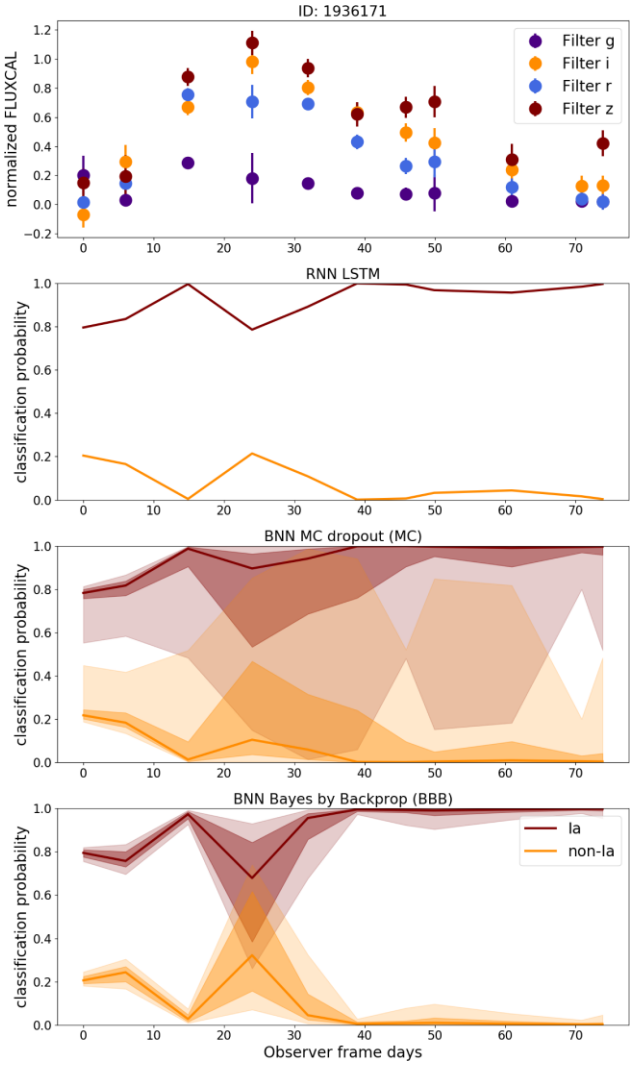

Figure 1. SNN classification for the DES-SN candidate DES17C2hqm at redshift 0.473 ± 0.001 using three different neural networks: baseline RNN, BNN MC dropout (MC), and BNN Bayes by Backprop (BBB). All methods were trained with the TRAIN-SIM simulation. Top row shows the SN candidate light curve from DES (normalized flux with *cosmo_quantile* method in band-passes *g, r, i, z*; time in Observer Frame days). Bottom rows shows the classification scores for each method (SN Ia: maroon, non-SN Ia: orange). Classification scores use all the data before a given date. The BNN methods provide classification uncertainties (shadowed regions show 68 and 95 per cent contours). Each BNN method provides different estimations, this is explored in Section 5.2. The large uncertainties in the classification probability represent the lack of confidence in this classification. For this example, uncertainties around days 20–30 are correlated with the lower SNR, while around days 50–60 that correlation is less straight forward to interpret and could be linked to the secondary peak visible in most filters.

their measurement uncertainties over time for accurate classification of time-domain candidates. Additional information such as host galaxy redshifts can be included to improve performance.

SNN includes different classification algorithms, such as LSTM[2] Recurrent Neural Networks (RNNs) and two approximations for Bayesian Neural Networks (BNNs). We show in Fig. 1 the classification probabilities from different methods for a given SN light curve.

[2]Long short-term memory (LSTM; Hochreiter & Schmidhuber 1997).

These probabilities can be used to select a sample by performing a threshold cut or by weighting the contribution of candidates by their classification score as in the BEAMS and BBC methods (Kunz et al. 2007; Kessler & Scolnic 2017; Vincenzi et al. 2022).

Light-curve simulations are used to train SNNto classify candidates into different classes. For cosmology, it can be trained to accurately classify SNe Ia versus other other kinds of transients. For time-domain astronomy, where brokers are designed to disentangle multiple types of transients, SNNcan classify subtypes of SNe or transients simultaneously.

Throughout this work we only perform a binary classification, i.e. a normal SN Ia or a non-Ia SN. Our results are expressed in the form of a prediction of the SN type by using a threshold on the obtained SN Ia probability, P, larger than 0.5.

3.3.1 SNN normalization schemes: cosmo and cosmo quantile

Since light-curve fluxes and uncertainties exhibit large variations, SNNsupports different input data (e.g. fluxes, flux-uncertainties, and time-steps) and normalization schemes (Möller & de Boissière 2019). In previous work, the default was the *global*[3] normalization. However, to avoid cosmological bias when using redshifts for classification, it is important to avoid using distance information encoded in the apparent magnitudes.

For classification using redshifts, we introduce two new normalization schemes in SNNthat ignore distance information: *cosmo* and *cosmo_quantile*.[4] In these schemes, for a given light curve, fluxes, and their respective uncertainties are normalized by the maximum light-curve flux in any filter (*cosmo*) or the 99th quantile of the flux distribution to avoid normalization using an outlier (*cosmo_quantile*). This normalizes the fluxes for each light curve to 1 or near 1, and retains colour and signal-to-noise information for the classification. The normalization of the time-step, given as an input to SNN, remains log transformed and displaced to zero as in the *global* normalization scheme.

To evaluate these new normalization schemes, we measure the classification accuracy of SN Ia versus non-SN Ia including redshift as an input using simulations from Möller & de Boissière (2019) since these were the simulations used to benchmark the SNNframework. We find that they slightly improve performance with accuracies of 99.33 ± 0.02 per cent for both *cosmo* and *cosmo_quantile* as compared to the 98.43 ± 0.08 per cent accuracy of the *global* normalization scheme using same data set, redshift information and default settings (seeds and hyper-parameters). In the following analysis, we will use only the *cosmo_quantile* norm since it has similar accuracy to *cosmo* for the simulations but is more robust against photometry outliers in real data.

3.4 SNN configuration for performance and robustness

We next study the performance of SNNwhen classifying SNe using photometry and host galaxy redshifts. We also characterize the classification robustness with respect to the training templates, and find the best set of hyperparameters for our DES-SNIa sample. We use the S-TRAIN-SIM simulations introduced in Section 3.2, for computational efficiency and to compare results with those of Möller & de Boissière (2019), to train a classification model. Our simulation was class-balanced (half normal SNe I and half non-Ia SNe) and randomly split in 80 per cent for training, 10 per cent for validation and 10 per cent for metrics evaluation. Uncertainties in the accuracy represent the standard deviation of predictions from five models obtained with different seeds.

Using the default configuration of SNNwe obtain a classification accuracy of 97.73 ± 0.04 per cent for the *cosmo_quantile* norm. While this accuracy is high, it is ∼ 1 per cent lower than the benchmark in Möller & de Boissière (2019) for a similar training set size. Since the SNNarchitecture has not been changed, we investigate if this can be attributed to the more complex and realistic DES-SN 5-yr simulations in Section 3.4.1. We then investigate whether a modified architecture can improve the classification model and thus its accuracy in Section 3.4.2. We highlight that SNN does not reach its peak performance when trained using the smaller S-TRAIN-SIMS. Thus, larger simulations are needed to improve the model performance.

Table 2. Classification accuracies for models trained by replacing a subset of templates from the original configuration in Section 3.2.

Changed template	Accuracy
JLA instead of extended SNIa model	97.96 ± 0.05
without peculiar SNe Ia	98.21 ± 0.01
J17 instead of V19 core-collapse model	98.06 ± 0.07
SPCC templates instead of V19 core-collapse model	98.59 ± 0.02

3.4.1 Templates impact on performance

Here, we study how the set of templates used to generate the training simulation impacts the metrics of our classification algorithm. We train different models using simulations that are similar in size (equivalent to S-TRAIN-SIM) but are generated by replacing a subset of templates from the original configuration. Obtained accuracies are shown in Table 2.

Models trained with SPCC and J17 templates obtain higher accuracies than those trained with V19 templates. This is consistent with the accuracy decrease of our present model when compared to that of Möller & de Boissière (2019). This is evidence of the more complex classification task with the updated simulations. We highlight that V19 uses a large variety of core-collapse templates with greater diversity than previous core-collapse models, J17 and SPCC. From these, SPCC has the fewest number of non-Ia templates and thus less diversity. SPCC templates were used in Möller & de Boissière (2019) simulations. The impact of changes like using the JLA SALT2 model is less. This shows that the complexity of the classification task increases largely with the updated and more diverse core-collapse SN population in the V19 templates and the inclusion of peculiar SNe Ia.

We thus attribute the decrease on accuracy to the more complex task of disentangling SNe Ia from core-collapse and peculiar SNe Ia generated with updated templates.

3.4.2 Hyperparameters

We investigate whether network hyperparameters could be modified to improve performance (for a list of available hyperparameters,

[3]Features, f, are log transformed and scaled. The log transform (f_l) uses the minimum value of the feature in all band-passes min(f) and a constant (ϵ) to centre the distribution at zero as follows: $f_l = \log(-\min(f) + f + \epsilon)$. Using the mean and standard deviation of the log transform (μ, $\sigma(f_l)$), standard scaling is applied: $\hat{f} = (f_l - \mu(f_l))/\sigma(f_l)$.

[4]Both normalization schemes are available at: https://github.com/supernnova/SuperNNova

Table 3. SNN Baseline Performance versus method on TRAIN-SIM without cuts. The chosen method in this work is the Ensemble (probability average) and is highlighted in bold.

Method	Balanced accuracy	Efficiency	Purity
	cosmo		
Single model	98.33 ± 0.01	98.65 ± 0.05	98.03 ± 0.06
Ensemble (target av.)	98.43 ± 0.02	98.81 ± 0.02	98.08 ± 0.02
Ensemble (prob. av.)	98.45 ± 0.01	98.80 ± 0.02	98.11 ± 0.02
	cosmo_quantile		
Single model	98.35 ± 0.01	98.68 ± 0.07	98.03 ± 0.05
Ensemble (target av.)	98.45 ± 0.00[5]	98.84 ± 0.02	98.09 ± 0.01
Ensemble (prob. av.)	$\mathbf{98.46 \pm 0.01}$	$\mathbf{98.83 \pm 0.03}$	$\mathbf{98.10 \pm 0.03}$

[5]We provide only two-significant figures. The uncertainties are negligible and less than 0.005.

Table 4. SNN baseline performance versus method on TEST-SIM with loose selection and JLA-like cuts.

Method	Balanced accuracy	Efficiency	Purity
	with loose selection cuts		
Single model	98.61 ± 0.03	99.61 ± 0.02	99.43 ± 0.02
Ensemble (prob. av.)	98.69 ± 0.01	99.68 ± 0.01	99.45 ± 0.005
	+ JLA-like cuts		
Single model	98.26 ± 0.06	99.81 ± 0.01	99.7 ± 0.01
Ensemble (prob. av.)	$\mathbf{98.36 \pm 0.01}$	$\mathbf{99.86 \pm 0.01}$	$\mathbf{99.71 \pm 0.005}$

see Möller & de Boissière 2019). We train our models using using 20 per cent of the S-TRAIN-SIM simulations (280 033 light curves). We modify: batch size (128, 512), dropout (0.05, 0.1, 0.2), bidirectional (True, False), hidden dimensions (32, 64, 128), number of layers (2,3,4), two learning policies (cyclic and non-cyclic), and different cyclic phases when using cyclic ([5, 10, 15], [20, 40, 60]). We find that the accuracy in different configurations varies up to ≈ 2 per cent. We find that deeper (3 or 4 layers) and wider networks (up to 64 hidden dimensions) result in the biggest changes to the accuracy. This reflects the increasing complexity of the classification task with updated SN templates. Our chosen configuration for S-TRAIN-SIM is: batch size 512, dropout 0.05, bidirectional network, 64 hidden dimensions, 4 layers, and non-cyclic learning policy. Using the whole S-TRAIN-SIM data set with this new configuration, the classification accuracy rises to 98.10 ± 0.06 per cent.

3.5 SNN trained models for DES-SN 5-yr analysis

In the following, we use SNNmodels trained with a larger data set to improve classification accuracy, TRAIN-SIM, and the best configuration of SNNfound in the previous section. We increase the batch size to 1024 for efficient resource allocation. The larger simulation and optimized hyperparameters provide a better classification accuracy with accuracies above 98 per cent as shown in Table 3. Accuracies are computed with a balanced test set, where half of the candidates are SNe Ia and half are non-Ia SNe.

To evaluate the accuracy, efficiency, and purity of our photometric samples, we estimate the performance of our models in the independent TEST-SIM. This simulation is not balanced and thus reflect the relative rate between SN types. We present performance metrics for different levels of selection cuts in Table 4. We highlight that we provide the balanced accuracy which shows that after the JLA-like cuts, the remaining non-Ia SNe are harder to disentangle. A thorough analysis on systematics linked to this classification method can be found in Vincenzi et al. (2022).

In this work, the traditional classification method is named 'single model'. This method represents classifications done using probabilities obtained from one SNN trained model with a single seed. In the following, we provide a mean value and uncertainty on the metric or classified sample of the 'single model' method by taking the probabilities obtained with 5 models trained with different seeds. These probabilities are then used to compute the mean and standard deviation of the metrics listed in Table 3.

3.5.1 *Ensemble methods*

For cosmology, we aim to have a classification method that is not highly sensitive to statistical fluctuations in the model and training data set. In ML, ensemble methods have been shown obtain more robust predictions (Dietterich 2000; Lakshminarayanan, Pritzel & Blundell 2017) and have been introduced for regression in astronomy (Carrasco Kind & Brunner 2014; Kim, Brunner & Carrasco Kind 2015). To produce ensemble classifications, predictions from multiple models are combined. This can be viewed as a mechanism of Bayesian marginalization (Wilson & Izmailov 2020; Izmailov et al. 2021) and an alternative to Bayesian Neural Networks using Variational Inference explored in Section 5.

We explore two possible ensemble methods: 'probability averaging' and 'target averaging'. Probability averaging uses the probability scores and averages them to select light curves that are above the 0.5 probability threshold of being SN Ia. The 'target average' method averages the predictions and selects the most common one. Uncertainties are computed using the standard deviation of the metric for three different sets of five models with different seeds.

We find that ensemble methods increase the accuracy and purity $\approx$ 0.1 per cent from just using one model prediction, or 'single model', as can be seen in Table 3. We find a 99.4 per cent overlap between photometrically selected Type Ia SNe using both the ensemble and single model methods. In the following, we will use the 'probability average' from different models as our ensemble method.

Each ensemble in this work is obtained using the predictions of 5 models trained with different seeds, also called an 'ensemble set'. To study the performance of ensemble methods, we compute metrics using the output of 3 ensemble sets, quoting their mean and standard deviation.

3.5.2 *Generalization*

In this Section, we verify the ability of our trained models to classify data generated using different simulation templates. This is called generalization and showcases the adaptation of our SNN models to new unseen data.

We evaluate the accuracy of our models when trained with simulations generated using SNe Ia, peculiar SNe and the V19 core-collapse templates but applied to simulations generated using other core-collapse templates such as J17 or SPCC. We observe a decrease of < 0.5 per cent in accuracy, which shows that our V19 trained models generalize well to other templates of core-collapse SNe.

We find that ensemble methods such as *probability average* reduces the loss in accuracy due to changes in the data by 0.2 per cent relative to the single model. This is expected as ensemble methods are usually more robust and thus generalize better than single models.

3.6 Bayesian Neural Networks (BNNs)

In scientific analyses using machine-learning outputs, it is important to evaluate the reliability of a model's predictions, expressed through uncertainties. Uncertainties can be divided into: *Aleatoric*, usually linked to measurement uncertainties (e.g. noise or other effects of data acquisition); *Epistemic* or model uncertainty, which encompasses uncertainties in the training set and NN architecture.

In this section, we introduce BNNs which are a promising method to provide uncertainties reflecting the model's confidence on the prediction.

To compute uncertainties, we obtain different classification probabilities for a given input and evaluate their variance. In NNs, this is equivalent to finding a posterior distribution of weights. Typically, this posterior distribution is intractable for deep neural networks, thus different methods have been developed to approximate it. A review on BNNs, approximation methods and their use in astronomy can be found in Charnock, Perreault-Levasseur & Lanusse (2020).

In this Section, we use two BNN implementations approximating the posterior distribution of weights: MC dropout (Gal & Ghahramani 2016) and Bayes by Backprop (Fortunato, Blundell & Vinyals 2017). MC dropout (MC in the following) provides a Bayesian interpretation by using the same dropout mask at the different NN layers including the recurrent ones (each time-step). Bayes by Backprop (BBB in the following) learns a posterior distribution of weights which can then be sampled. Both methods have been previously implemented and tested on simulations in SNN (Möller & de Boissière 2019).

3.6.1 *BNN classification probabilities and uncertainties*

For both methods, to obtain the classification probability distribution, we sample the predictions from our BNN 50 times. This sampling number is also known as the number of *inference samples*, n_s. In the following we compute the classification probability, P_i for a given light curve, $\mathbf{x}_i$ as the mean of sampled probabilities:

$$P_i = \frac{1}{n_s}\left(\sum_{j=1}^{n_s} p_j(\mathbf{x}_i)\right), \tag{1}$$

where $j \in [1, n_s]$ is the index of inference samples, $p_j(\mathbf{x_i})$ is the j^{th} sample of the classification probability distribution for the light-curve $\mathbf{x}_i$.

We compute the classification probability uncertainty for a given light-curve $\mathbf{x}_i$ as the standard deviation of sampled probabilities:

$$\widehat{\sigma_i} = \frac{1}{n_s}\sqrt{\sum_{j=1}^{n_s}\left(p_j(\mathbf{x}_i) - P_i\right)^2}, \tag{2}$$

where $j \in [1, n_s]$ is the index of inference samples, $p_j(\mathbf{x_i})$ is a classification probability for the given light-curve $\mathbf{x}_i$ for each inference sample j, and P_i is given by equation (1).

3.6.2 *BNN trained models*

Using the TRAIN-SIM simulations we train the two Bayesian models, MC and BBB, for light-curve classification with host galaxy redshifts. Both methods obtain high classification accuracies for the ensemble probability average method, 98.33 ± 0.01 and 98.11 ± 0.01 for MC dropout and BBB, respectively. Balanced accuracies are slightly lower than the ensemble method in Table 3. These may be improved by adjusting of the hyperparameters. We choose to keep the current configuration and focus on the behaviour of the classification uncertainties.

Table 5. Performance metrics of BNNs evaluated using TEST-SIM simulations with JLA-like cuts. These simulations are indicative of the expected purity and efficiency of our photometrically classified samples.

Method	Accuracy	Efficiency	Purity
MC with JLA-like cuts			
Single model	98.01 ± 0.03	98.41 ± 0.03	97.63 ± 0.07
Ensemble (prob. av.)	98.11 ± 0.01	98.51 ± 0.06	97.73 ± 0.05
BBB with JLA-like cuts			
Single model	98.01 ± 0.03	98.41 ± 0.03	97.63 ± 0.07
Ensemble (prob. av.)	98.11 ± 0.01	98.51 ± 0.06	97.73 ± 0.05

Traditionally, BNNs are not used in ensembles, combining predictions by different models. To do so, ideally, the probability distributions for each model in the ensemble set should be concatenated into a 'joint probability distribution'. Then, the ensemble classification probability would be computed using equation (1) sampling n_s times the 'joint probability distribution'. However, this can be computationally expensive. Using TEST-SIM simulations, we find that averaging the mean probability obtained for each model in the ensemble set is a close approximation of the one obtained using 'joint probability distribution'. We find that the differences between probabilities using the approximation and the 'joint distribution' are centred at 0.00 ± 0.01 and accuracies change by less than 0.1 per cent. We use this approximation in the following for computational efficiency.

We also test approximating ensemble uncertainties as the sum of uncertainties from each model in the ensemble set assuming the covariance between models is zero. We find on average that the uncertainties obtained with this approximation and from the 'joint probability distribution' are similar. However, we note that the approximation for the BBB method has a larger dispersion than the one for the MC method. We will evaluate the potential use of BNN classification uncertainties in Section 6.2.

We use TEST-SIM to evaluate the expected metrics for our photometrically classified samples with JLA-like cuts in Table 5. The samples obtained with BNNs have less than 3 per cent contamination but that is higher than our Baseline DES-SNIa samples with JLA-like cuts. BNN performance could be eventually be improved with a different network configuration and initialization. However, for comparison we keep this architecture for the analysis in Section 5.

4 DES-SN 5-YR PHOTOMETRICALLY CLASSIFIED SNE IA

In this Section, we photometrically classify DES-SN 5-yr candidates with host spectroscopic redshifts using our baseline RNN trained in Section 3.5.

First, we classify candidates that pass loose cuts using SNN trained with host galaxy redshifts in Section 4.1. We further constrain the sample using JLA-like cuts and visual inspection in Section 4.2. We discuss possible contamination of this sample in Section 4.3 and its classification efficiency in Section 4.4. We summarize the properties of the baseline photometrically classified SN Ia sample with JLA-like cuts in Section 4.5.

4.1 Photometric classification

We use our baseline RNN model to select photometrically classified SNe Ia. We show the number of selected light curves in

Table 6. DES-SNIa photometric samples with different selection cuts. In the last row, we define the **Baseline DES-SNIa sample** using a single ensemble set probability threshold. Columns indicate: the number of photometrically selected SNe Ia and the number of spectroscopically classified SNe Ia contained in that sample.

Method	Loose selection cuts		+ JLA-like cuts	
	photo Ia	spec Ia	photo Ia	spec Ia
Single model	1861^{+17}_{-13}	353^{+3}_{-3}	1478^{+10}_{-11}	320^{+3}_{-2}
Ensemble (prob. av.)	1867^{+3}_{-4}	354^{+1}_{-0}	1482^{+2}_{-2}	321^{+1}_{-0}
Baseline DES-SNIa sample	**1863**	**354**	**1484**	**321**

Table 6 and their overlap with spectroscopic SN samples defined in Section 2.

As shown in Sections 3.5.1 and 3.5.2, ensemble methods provide more robust predictions than single model methods. We select our Baseline SNe Ia sample using the 'probability average' method and the *cosmo_quantile* norm. This normalization is more robust towards photometry outliers present in our analysis. We note that the overlap between *cosmo* and *cosmo_quantile* probability average sample is larger than 98 per cent and between *cosmo_quantile* probability average and single model samples is larger than 99 per cent.

Our Baseline DES-SNIa sample contains 1863 photometrically identified SNe Ia passing loose selection cuts. In this sample, 12 spectroscopically classified SNe Ia are not selected, representing less than 1 per cent of the photometric sample. We do not find a particular redshift or SALT2 parameter preference for these lost SNe Ia. Visual inspection reveals some light curves have variable quality photometry which could contribute to the misclassification.

The baseline sample with loose selection cuts can be used to study astrophysical properties of SNe Ia like correlations with their host galaxies, diversity, and rates. In the following, we further constrain this sample with cosmology-grade cuts as in Vincenzi et al. (2022).

4.2 Cuts towards a cosmology sample (JLA-like)

We further constrain our sample by applying selection cuts based on SALT2 light-curve fits and redshift quality.

First, we implement additional requirements on the fitted SALT2 parameters of the photometrically selected SNe Ia. As in Vincenzi et al. (2022), we implement the JLA-like SALT2 cuts from the Joint Light-curve Analysis (Betoule et al. 2014) introduced in Section 2.3.2. Secondly, we select only candidates which have a high-precision spectroscopic redshift. We eliminate those candidates that have redshifts provided by PRIMUS since the spectra are of lower resolution, more prone to catastrophic failures and not high-quality enough for cosmology analysis.

The results of these cuts in the photometrically selected samples are shown in Table 6. We highlight that the JLA-like cuts reduce the scatter in the number of SNe, as can be seen by the reduced standard deviation in the table when compared to the sample without JLA-like cuts. We obtain a Baseline DES-SNIa sample with JLA-like cuts of 1484 photometrically classified SNe Ia. The missing spectroscopic SNe Ia are found to be redder in average and at all redshifts with a median around 0.5.

A summary of the selection criteria used to obtain this sample can be found in Table 7. General properties of these samples are further studied in Section 4.5.

4.3 Contamination

As shown in Vincenzi et al. (2022) and in Table 4 contamination from core-collapse and peculiar SNe in a SNN classified sample with quality cuts is expected to be less than 1 per cent. This estimate was obtained using SN simulations containing various types of core-collapse and peculiar SNe. We inspect the Baseline DES-SNIa sample with JLA-like cuts obtained in the previous section and do not find any spectroscopically identified core-collapse or peculiar SNe. We note that spectroscopic samples are not complete and DES-SN follow-up preferentially targeted suspected Type Ia SNe.

In this section, we explore a different type of potential contaminant, 'out-of-distribution' candidates such as AGNs and other unknown transients. These candidates can be erroneously classified since they are not present in the simulated training sample and thus we do not know how SNNclassifies them.

We find no spectroscopically identified AGN, SLSNe, or other SN spectral types in our Baseline DES-SNIa sample but 5 candidates with host spectra showing AGN features. We find that DES16E2nb, DES16X1ext, DES13X3dbe are displaced by more than 1 arcsec from the centre of the galaxy (additionally DES16E2nb is a spectroscopic Type Ia SN) and the other two candidates are displaced between 0.5 and 1 arcsec. At these separations, the light curves from these candidates are not dominated by the AGNs which we confirm by inspection of the light curves. Therefore we keep these photometrically selected SNe Ia in our Baseline DES-SNIa sample.

We also perform visual inspection of the light-curves in the Baseline DES sample. We find three candidates that can be visually tagged as multiseason visually: DES16E2nb a spectroscopic SN Ia with close by AGN, DES16C3nd two SN Ia in a galaxy (Scolnic et al. 2020), DES14E2rpm a spectroscopic SN Ia with a fake SN inserted at the same coordinates (fakes were inserted to evaluate the detection efficiency in DES-SN images, see Brout et al. 2019).We keep all these candidates since they are real supernovae with fake or other SN light curves that do not overlap.

Photometrically classified Type Ia SNe samples are expected to have some level of contamination from core-collapse and peculiar SNe and possibly by other transients. For the Baseline DES-SNIa sample in this work we find no clear evidence of contamination from core-collapse and peculiar SNe or long-term variables such as AGNs.

4.4 Classification efficiency

Traditionally, in cosmology analyses using spectroscopically classified SNe samples, modelling selection effects is crucial to estimate biases and systematic uncertainties.

Selection effects arise from a combination of SN detection and other effects. They are usually modelled as an efficiency with respect of an observed magnitude. For host galaxy selection, Vincenzi et al. (2020) use the host galaxy r band magnitude, $m_r^{\rm host}$. For spectroscopic classification, Smith et al. (2020) and Kessler et al. (2019b) use the modelled supernova peak magnitude in the i band, $i_{\rm peak}$ computed from the best-fitting SALT2.

To determine if there is a selection efficiency decrease due to photometric classification, we inspect the differences between the peak observed magnitude in the i band of our Baseline DES-SNIa sample compared to simulated SNe Ia in DES-SN 5-yr in Fig. 2. Our Baseline DES-SNIa photometric sample follows the expected SN Ia peak magnitude distribution from simulations but we find an excess on the maximum magnitude with a reduced $\chi^2_\nu = 2.1$. We do not find evidence for additional selection efficiency effects from the photometric classification procedure.

Table 7. Effect of the selection cuts on the candidate sample. We show results for the shallow and deep fields, as well as the total number. Note that some events belong to both shallow and deep fields due to field overlap. Columns show the cut, the number of selected candidates, the number of spectroscopic SN Ia in the sample and the Section where the sample is described.

Cut	Shallow		Deep		Total		
	Selected	spec Ia	Selected	spec Ia	Selected	spec Ia	Section
DES-SN 5-yr candidate sample	29 203	415	7500	93	31 636	415	2.1
Multi-season	13 868	405	4428	88	14 070	405	2.2
Redshifts in $0.05 < z < 1.3$	6556	401	1812	85	6590	401	2.3.1
SALT2 loose selection	2380	366	698	77	2381	366	2.3.1
RNN>0.5 (**Baseline DES-SNIa**)	1863	354	502	76	1863	354	4.1
JLA-like (**Baseline DES-SNIa JLA**)	1484	321	408	73	1484	321	4.2

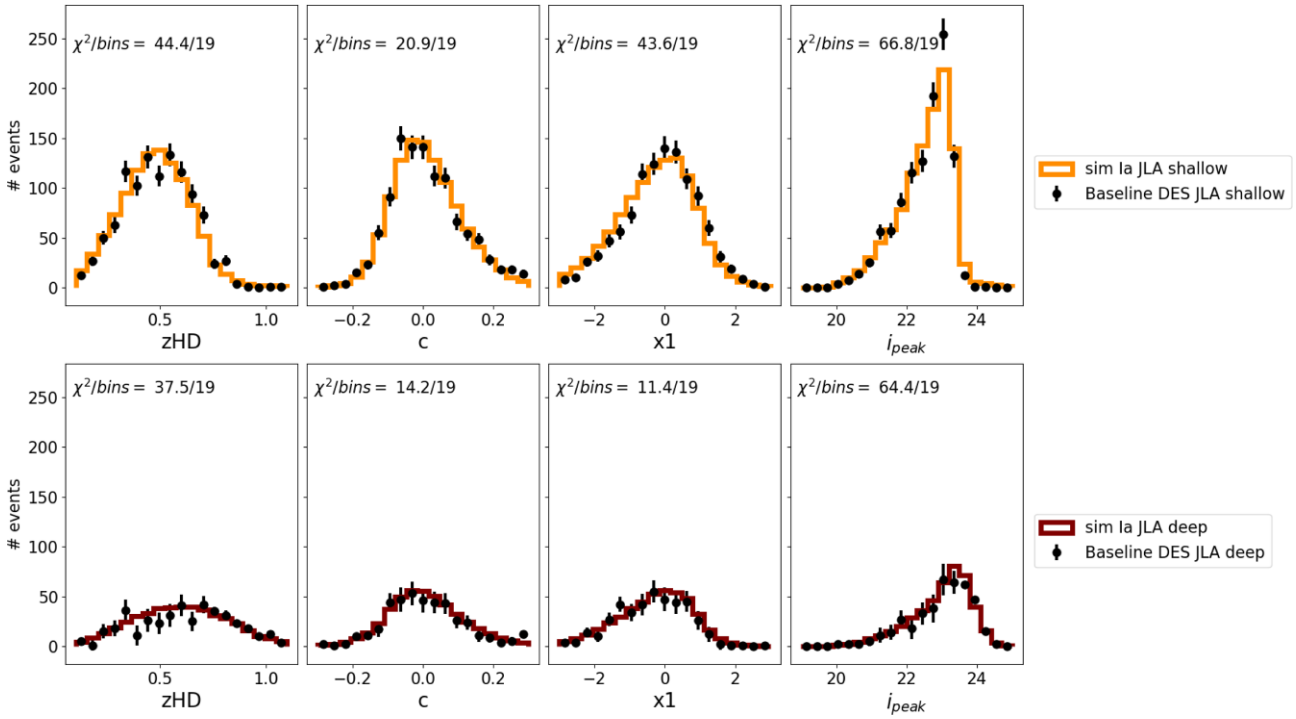

Figure 2. Distributions of redshift, SALT2 x1, SALT2 c, and peak magnitude in i band $i_{\rm peak}$ for our Baseline DES-SNIa sample from Section 4 for the shallow (yellow) and deep (maroon) fields. We show one simulated realization of DES-SN 5-yr sample. Poisson uncertainties are assumed. Both the simulation and data pass JLA-like cuts. The goodness of fit for each histogram is shown as the χ^2/number of bins on each plot.

4.5 Colour and stretch evolution

We study the properties of the Baseline DES-SNIa sample with JLA-like cuts and compare it to that expected from realistic simulations. In Section 4.4, we found that the effects of classification efficiency are negligible, thus we do not correct for this efficiency and use simulations including only detection and host galaxy redshift efficiency introduced in Section 3.2.

Fig. 2 shows the redshift *zHD* and SALT2 fitted colour c, stretch x_1 and $i_{\rm peak}$ distributions for the DES-SNIa 5-yr photometric sample classified using host galaxy redshifts. Fig. 2 also shows one realization of a DES-SN 5-yr simulated SNe Ia. Uncertainties are calculated as the square-root of the number of candidates per bin. There is decent agreement between the simulation and data, although the reduced χ^2_ν are somewhat larger than expected from statistical fluctuations.

In Fig. 3, we show the redshift evolution of our sample's colour and stretch. Our baseline sample matches the trends expected from the simulation. Although there are some slight differences outside the 68 per cent simulation contour (equivalent to 1σ for a Gaussian distribution) in particular for the shallow fields.

These differences might result from the small number of candidates (the last two redshift bins have only 24 and 16 SNe Ia), unaccounted classification contamination, unaccounted selection effects or whether there is redshift evolution in the intrinsic SN population (Scolnic & Kessler 2016; Nicolas et al. 2021; Popovic et al. 2021) or the effect of dust needs to be introduced (Jha, Riess & Kirshner 2007; Mandel, Narayan & Kirshner 2011; Mandel et al. 2017; Brout & Scolnic 2021). The optimization of the simulation and systematics studies is outside the scope of this work.

We now turn to select other photometric samples using the novel Bayesian Neural Networks and explore their possible use.

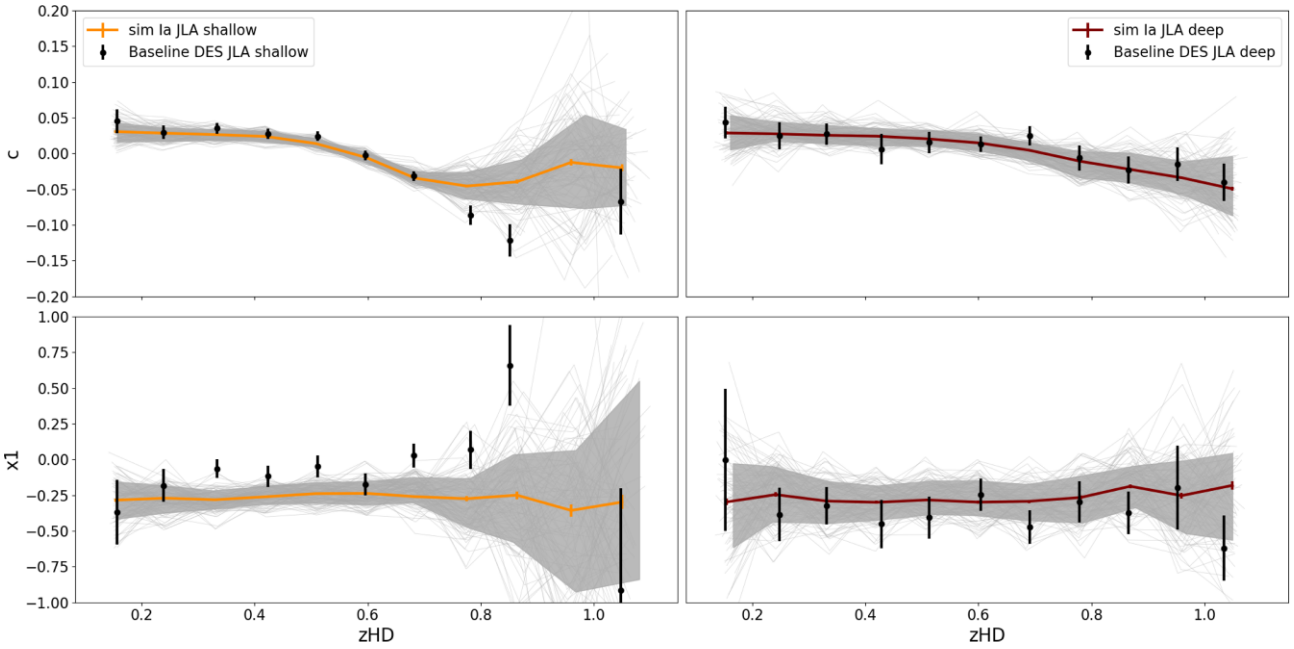

Figure 3. Redshift dependence of SALT2 c and x_1 for he Baseline DES-SNIa photometric sample and simulated SNe Ia for shallow (yellow, left) and deep (right, maroon) fields using the DES-SN host galaxy spectroscopic efficiency (Vincenzi et al. 2020) both with JLA-like cuts. For the simulation, orange lines are rolling averages of the measured parameters, in grey 150 realizations of SNe Ia in the DES-SN 5-yr survey and in solid grey the area covered by the 68 per cent of these realizations. The mean and the standard deviation are shown for data using black markers.

5 PHOTOMETRICALLY CLASSIFIED SNE IA WITH BAYESIAN NEURAL NETWORKS

In this section we explore the use of BNNs for classification. While the accuracy of these Networks is equivalent to the baseline RNN used in Section 4, BNNs also provide classification uncertainties.

We first obtain photometric samples using two BNN schemes (MC and BBB, Section 5.1). We then evaluate the classification uncertainties from BNNs (Section 5.2), and summarize our findings (Section 5.3).

5.1 BNN photometric sample

We apply our BNN trained models to candidates passing loose and JLA-like cuts introduced Sections 2.3.1 and 2.3.2. This candidate sample contains 1701 light curves that are then photometrically classified, as shown in Table 8.

Using BNN probabilities, the average probability ensemble method and a threshold of P larger than 0.5, we obtain about 3 per cent more candidates than our Baseline DES-SNIa sample with JLA-like cuts in Table 6 for both BNN methods. The additional BNN selected supernovae, 52 MC and 51 BBB, have distributions of colour, stretch, and redshifts that are representative of the Baseline DES-SNIa sample selected using the RNN models (Section 4). We find that 1 and 6 SNe Ia in the Baseline DES-SNIa sample are not selected by MC and BBB methods. These missing SNe Ia have red colours and are at median redshifts close to 0.5. The BNN samples are thus probing a similar parameter space to the Baseline DES-SNIa sample.

As in the previous sample, we find no spectroscopically identified AGNs, SLSNe, or other SN spectral types in our BNN photometric sample. We find the same 5 candidates with nearby spectra showing AGN features which are kept due to their large enough separation >0.5 arcsec, with the AGNs. In a cosmological sample, however, these candidates will be eliminated due to possible issues with the measured photometry.

5.2 BNN uncertainties

In this Section we try to interpret which types of uncertainties are captured in the outputs of the BNN model: *aleatoric* or *epistemic*. BNNs provide classification probability distributions that a priori indicate a confidence level on the prediction. These uncertainties are shown in Fig. 1 for each classification step. Here, we only evaluate the final uncertainty (final time-step) for each event.

In Fig. 4, we show the distribution of classification uncertainties for different samples. We compare the uncertainties derived from the data and from simulations. For most samples, the simulation and data uncertainty distributions are similar. This indicates that the simulations and data resemble closely after JLA-like cuts. However, a large difference is found where there is no selection cut which is further explored in Section 6.2.

Both BNN methods provide different order of magnitude of uncertainties estimates and distribution of mean uncertainties (e.g. BBB is more clustered in low uncertainty regions), possibly due to initialization parameters or intrinsic properties of the method. Accounting for those differences is not straight-forward, see Möller & de Boissière (2019) for a discussion on this topic.

We compare BNN uncertainties as a function of light-curves properties in Fig. 5. We find that MC dropout and BBB exhibit different behaviours for both data and simulations.

We find both indications in favour (+) and against (−) interpretation of classification uncertainties as a particular type:

a. *aleatoric uncertainty*: linked to measurement uncertainties (+) classification uncertainties are correlated to SNR in data. Bright candidates and those with higher quality light curves have on average smaller classification uncertainties for both BNNs.

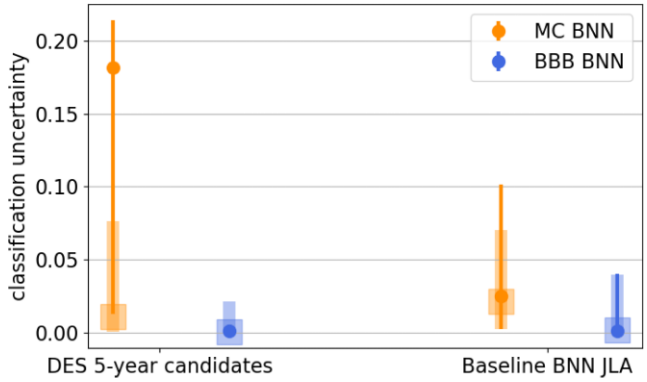

Figure 4. Classification uncertainties obtained for BNN ensemble models. Columns indicate which sample is used. For each event in a given sample, we obtain their classification uncertainties from the two BNN methods, MC and BBB (orange and blue, respectively). We show median uncertainties for data in circles for: all DES-SN 5-yr data (no selection cuts), and Baseline BNN SNIa samples with JLA-like cuts. For comparison, we show in squares the median uncertainties obtained for the whole simulation (first column) and simulated photometric samples with JLA-like cuts (second column). For both the data and simulations, we show as errorbars the extent of the 68 per cent of the distribution. The different behaviour of simulated MC uncertainties and that of DES-SN 5-yr candidate sample is further studied in Fig. 7.

(−) this correlation is not seen in the simulations for any of the BNNs.

b. *epistemic uncertainty*: linked to training sets or model

(+) Large uncertainties are more prevalent in classification probabilities far from 1 (high probability of being a SN Ia) and 0 (low probability of being SN Ia) for both simulations and DES-SN 5-yr data.

(−) candidates that fulfil selection cuts should more closely resemble simulated SNe Ia, thus it is puzzling the increase on median uncertainty when applying cuts in particular for the MC method (see Fig. 4).

These various behaviour highlights the challenges on quantifying uncertainties in complex problems such as astronomical data classification. In Appendix A, we explore further correlations between classification uncertainties and SALT2 fit light-curve properties.

We continue exploring the interpretability of the BNNs uncertainties by adding a threshold on the uncertainties for SNIa sample selection, as in Möller & de Boissière (2019) and more recently in Butter et al. (2022). We note that establishing a threshold for uncertainties is not straight-forward. While the whole probability distribution has a calibration that can be verified using diagnostic as reliability diagrams (DeGroot & Fienberg 1983; Möller & de Boissière 2019), the probability uncertainties do not. We chose to eliminate candidates with the highest uncertainties (eliminating candidates that are outside of 99 percentile of the uncertainty distribution). This cut rejects candidates that were in the RNN sample: 12 for the MC model and 45 for BBB. These candidates are not found to be distributed preferentially in a c, x_1, or redshift. We visually inspect these light-curves and found that a large proportion have photometry that are outliers.

5.3 BNN photometric sample contribution

The SNIa samples obtained using BNN methods are found to be similar to the one provided by our Baseline DES-SNIa sample in Section 4. We evaluate BNN uncertainties and show that they are consistent between simulations and data in average after JLA-like cuts, showing a good agreement between data and simulation predictions. However, BNN uncertainties are difficult to interpret and assess quantitatively (e.g. assigning an uncertainty threshold).

We find that uncertainties exhibit different behaviours in the two BNN methods and between data and simulations. While the higher uncertainties in the MC BNN method for the data could point towards the presence of out-of-distribution candidates, the evidence is not conclusive and is not seen in the BBB method. We will further explore the possible contribution of BNNs in photometric classification without any selection cuts in Section 6.2.

Cuts on uncertainty values potentially improve our photometric SNIa samples by rejecting candidates with photometry that contains outliers. These is a promising avenue shown to improve the quality of samples, both in quality of the data and rejection of out-of-distribution events, in previous work using simulations Möller & de Boissière (2019) and more recently with astronomical data in Butter et al. (2022).

6 FROM DES TO RUBIN OBSERVATORY LSST

For the LSST survey, where up 10^7 SNe will be detected over 10 yr, photometric classification will become increasingly important.

In this work, we have presented different methods for photometric classification with redshift information. We compare the samples obtained with these different methods in Section 6.1 and explore possible applications of Bayesian Neural Networks in future surveys, such as LSST, in Section 6.2.

6.1 DES-SNIa photometric samples

The DES-SN 5-yr data contains thousands of potential SNe Ia. We show in Table 7 the different steps used in this work to obtain our Baseline DES-SNIa JLA sample from the DES-SN 5-yr candidate sample. Cuts applied before photometric classification reduce the candidate sample by 90 per cent. Photometric classification and JLA-like cuts refine the sample with a small 20 per cent reduction. While this reduction is small, it reduces contamination from $\sim$ 10 per cent to below 1.4 per cent, as shown in Vincenzi et al. (2022) and in Section 4.

In addition to our Baseline DES-SNIa sample classified using RNN probabilities, we have explored identifying samples with Bayesian Neural Networks. We compare these samples with with the preliminary DES-SN 5-yr spectroscopically classified SNe Ia sample in Fig. 6. As expected, we find that photometric samples using RNNs or BNNs provide larger numbers of SNe Ia than the spectroscopic sample, probing a larger parameter space. We do not find a substantial difference in the parameter distributions between different photometric classification methods.

We highlight that the photometric samples peak at fainter magnitudes and higher redshifts than the preliminary DES-SN 5-year spectroscopic SNe Ia sample.This has the potential to reduce selection biases and opens the possibility of stronger statistical analyses with the large numbers of SNe Ia. This will also be true for the immense SN samples obtained with LSST.

6.2 Bayesian Neural Networks as a proxy

Introduced as a promising method to quantify model uncertainties, BNNs have not yet been widely used in classification tasks. In Section 5, we have shown the difficulties for uncertainty interpretation given the different uncertainty values for the BNN methods.

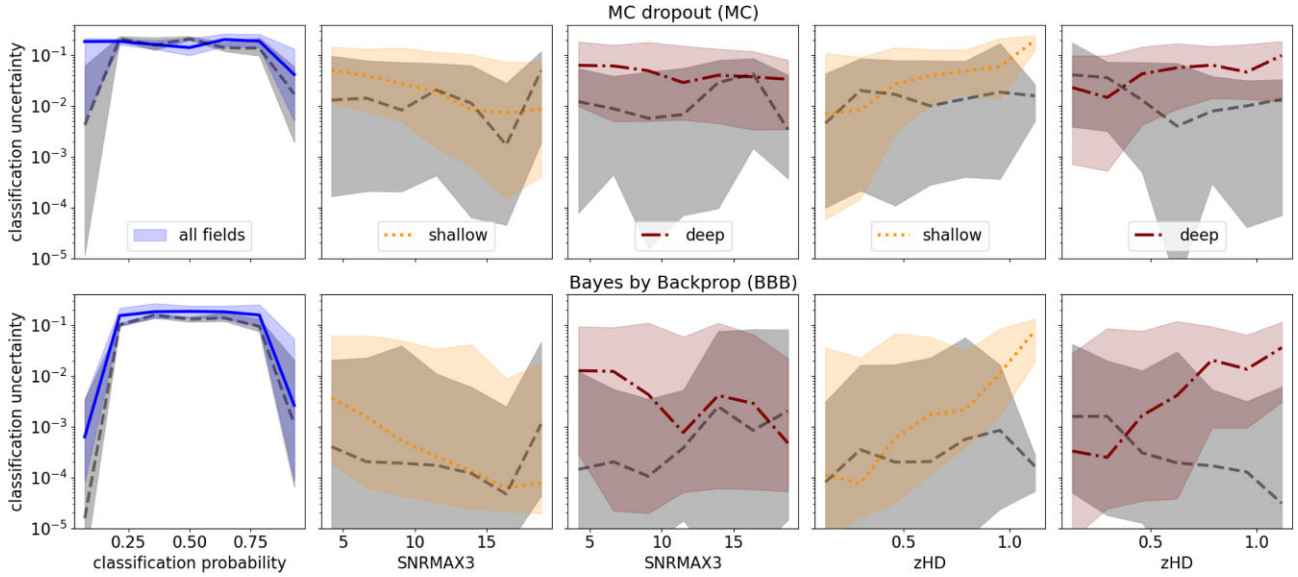

Figure 5. Distribution of classification uncertainty for Baseline MC dropout (upper row) and Bayes by Backprop technique (lower row). We show uncertainties as a function of classification probability for all fields (left), SNR of the third brightest point in the light curve (SNRMAX3, columns 2 and 3), and redshift (zHD, columns 4 and 5). Coloured lines show the median of the data with solid blue representing all fields, dotted yellow representing shallow fields, and dotted–dashed red representing the deep fields. Simulations are shown by the grey-dashed lines. Shaded regions show the 68 per cent percentile.

Table 8. Photometric classification of light-curves with Bayesian Neural Networks. Columns indicate: the number of photometrically selected events and the number of spectroscopic SNe Ia contained in that sample. We show these samples with JLA-like SALT2 cuts as in Section 4.2 and when adding a cut in the BNN classification uncertainty.

Method	+ JLA-like photo Ia	+ JLA-like spec Ia	+JLA-like + unc photo Ia	+JLA-like + unc spec Ia
		MC dropout		
Single model	1532^{+7}_{-4}	335^{+1}_{-1}	1513^{+6}_{-3}	333^{+0}_{-0}
Ensemble (prob. av.)	1535^{+3}_{-2}	336^{+0}_{-0}	1520^{+2}_{-1}	333^{+0}_{-0}
Baseline MC sample	1535	336	1520	333
		BBB		
Single model	1526^{+8}_{-6}	334^{+1}_{-0}	1487^{+5}_{-2}	328^{+2}_{-0}
Ensemble (prob. av.)	1528^{+1}_{-1}	335^{+1}_{-0}	1483^{+0}_{-0}	324^{+1}_{-0}
Baseline BBB sample	1529	336	1483	324

However, a potential use could be rejecting candidates with large uncertainties, as they sometimes have light curves with photometry outliers.

Here, we explore other possible uses of BNN uncertainties, using samples that have not been constrained with selection cuts. We aim to answer two questions: (i) can BNN uncertainties be used as an indicator of the representativity of the training set for a given data set? (ii) can BNN uncertainties replace selection cuts? We address these questions in Sections 6.2.1 and 6.2.2, respectively. The former could be useful to choose the set of SED templates to simulate a survey. As some selection cuts require feature extraction, the latter could be valuable to avoid this time-consuming process by using instead classification uncertainties from non-parametric classifiers as SNN.

6.2.1 BNNs uncertainties versus simulation representativity

First, we use simulations to assess the expected behaviour of uncertainties when training sets are not representative of the testing data.

We examine how the uncertainties change when using the trained model in Section 3.6.2 and applied to individual simulations with normal Type Ia supernovae and core-collapse SNe generated with the V19, SPCC and J17 templates. We expect that the trained model is representative of the V19 simulation. This will not be true for J17 and SPCC.

We find that both the single seed and ensemble methods have accuracies which decrease for J17 and SPCC simulations by ≈ 0.5 per cent for both types of BNNs. We see an increase in the *mean model uncertainty* on classified light curves generated with J17 and SPCC, however this change is within uncertainties. For both BNNs we find a longer and more significant tail for the uncertainty distributions when classifying J17 and SPCC simulations (ending at ∼0.4–0.43 compared to ∼0.35 for V19).

Next, we compare uncertainties when classifying DES-SN 5-yr data with independent BNN models trained with the V19, J17, and SPCC simulations. We find that the *mean model uncertainty* increases for SPCC and J17 classification models for MC dropout but not for BBB SPCC model but again within uncertainties. The tail of the uncertainties varies between ∼0.40 and 0.47 for all classification models. We see a longer tail for the uncertainty distributions for BBB but not for MC SPCC classification.

In summary, we do not find strong evidence of BNN uncertainties being sensitive to models trained with different core-collapse templates. There is a small but inconclusive tendency to increase uncertainties for J17 and SPCC in simulations. While these templates are different, the changes may be too small to be captured by BNN uncertainties.

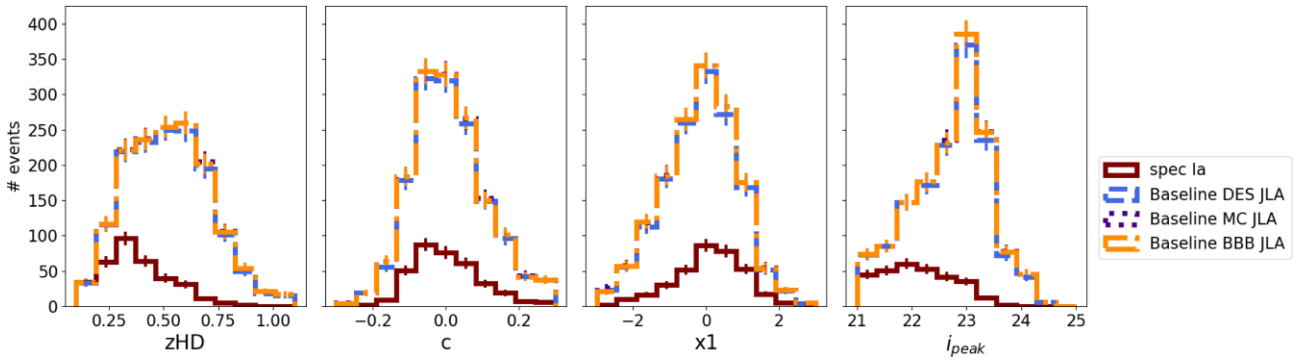

Figure 6. Distributions of redshift, SALT2 c, x1 and peak magnitude in the *i* band i_{peak}, for the samples with JLA cuts: preliminary DES-SN 5-yr SNIa spectroscopically classified (maroon), Baseline DES-SNIa (RNN) (blue) and Baseline MC and BBB SNIa samples, purple and orange, respectively. We note that the MC and BBB samples distributions almost completely overlap one other.

6.2.2 BNN uncertainties as a proxy for selection cuts?

We further study the distribution of classification uncertainties for samples selected with different cuts.

First, we check the behaviour of uncertainties with simulations. Uncertainties are distributed with a peak at low values and a decreasing long tail. We find that as the sample is refined through cuts in redshift, SALT2 convergence, and others, the maximum uncertainty is reduced. For example, if the simulated sample passes loose selection cuts and then a JLA-like cut is applied, the maximum uncertainty in the distribution reduces from 0.37 to 0.26 in MC dropout and from 0.34 to 0.25 in BBB. We do not find a significant change in the median distribution since it is dominated by small uncertainty values.

For the DES-SN 5-yr data we show the distribution of classification uncertainties in Fig. 7 with different selection cuts (see Section 2.3.1). As selection cuts are applied, the maximum uncertainties reduces for both methods as in simulations.

We highlight an interesting behaviour seen for MC dropout classification uncertainties. We find that this method assigns high uncertainties to candidates that do not have a secured redshift and candidates that are filtered with the multiseason cut. While the model was trained to use host galaxy redshifts, it can provide a classification for objects using a default value provided, here an assigned redshift of −9. While these candidates are clearly outliers (the redshift provided for classification is −9) and can be eliminated using simple cuts, this could indicate that MC dropout uncertainties are indicative of out-of-distribution candidates. Importantly, many of these high-uncertainty candidates are classified with probability larger than 0.5 which, without selection cuts, would end up in our photometric sample if no selection cuts were applied. We do not see this behaviour in the BBB model.

The multiseason veto and redshift availability cut effectively eliminates the light curves producing the high-uncertainty peak for MC dropout. After these cuts, the most impactful cut for higher uncertainties is linked to the SALT and JLA quality cuts. This is not surprising since these cuts restrict the SN properties range to the ones for normal SNe Ia.

In summary, we find that BNN methods behave differently when classifying out-of-distribution candidates defined as light curves without redshift. Interestingly, the high-uncertainty peak found for the MC dropout method in Fig. 7 reflects a possible interpretability of these uncertainties. This interpretability could help to quickly identify the presence of anomalies in the data set which were not in the training sets of the model.

For current surveys, our candidate samples are small enough to easily identify out-of-distribution events using feature distributions. However, for future surveys such as Rubin LSST this may prove difficult given the expected detection of 10 million transient candidates per night. Here we find that BNN uncertainties from MC dropout scheme can provide an indication whether there are out-of-distribution events in a given candidate sample and further selection cuts may be required.

7 CONCLUSIONS

In this work we train Type Ia versus non Ia classification models using large realistic DES-like simulations and apply them to DES-SN 5-yr data.

We introduce pre-processing of DES-SN light curves for accurate photometric classification. This includes selection of light curve time-span, photometry quality cuts, and selection cuts to limit out-of-distribution candidates that are not included in the training set (e.g. AGNs).

We present samples classified with host galaxy redshifts using SNNRecurrent Neural Networks and explore the use of Bayesian Neural Networks. We introduce the use of ensemble predictions for SN classification. We find that selecting SNe using an ensemble of models is more robust and stable than any single model.

Using host galaxy spectroscopic redshifts, we select a Baseline DES-SNIa sample of 1863 photometrically identified Type Ia SNe. This sample can be used for astrophysical studies of the properties of SNe Ia and their environments. For cosmology, we apply JLA-like cuts and select 1484 photometrically classified SNe Ia. This sample is more than three times larger than the DES-SN 5-yr spectroscopically confirmed SN Ia sample and covers a larger redshift range. Most of the spectroscopically identified SNe Ia in DES-SN are included in this photometric sample. These 1484 photometrically identified SNe Ia are currently the largest single-survey high-quality SN Ia sample and is being used for studies such as rates and SNe Ia host-galaxy properties.

We find that the properties of the SNe Ia in our Baseline DES-Ia sample are reproduced in the simulations. We anticipate that with further refinements (improved host galaxy libraries and more accurate dust models), the agreement between the simulations and the data will improve.

Additionally, we explore the use of uncertainties provided by Bayesian Neural Networks for identifying out-of-distribution candidates and defining representative training sets. We highlight some

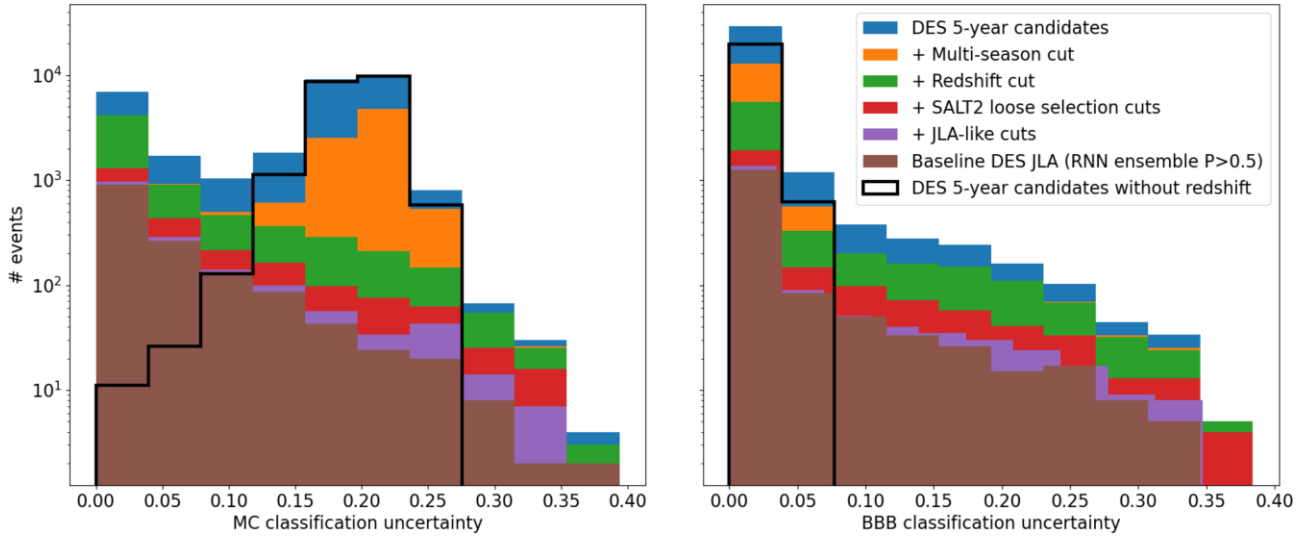

Figure 7. Uncertainties obtained with the two BNN methods (MC and BBB) for the DES-SN 5-yr candidate sample, through different selection cuts: multiseason filtering, redshifts, SALT2 convergence and JLA-like cuts; and our photometrically identified sample (filled histograms). We show the number of events in the *y*-axis in log scale. MC dropout uncertainties seem to identify those out-of-distribution candidates that have no redshift information (black line) or are filtered multiseason events. This secondary peak drives the mean uncertainty behaviour for MC dropout in Fig. 4.

of the BNN pitfalls and the difficulty of comparing classification uncertainties between variational inference methods. We find that the MC dropout BNN provides potentially interpretable uncertainties for out-of-distribution event detection and improving the photometric sample. This work is the first known application of two BNN methods on real astrophysical data for classification tasks.

This work is part of the DES-SN 5-yr cosmology analysis. We have optimized simulations, the SNNarchitecture, as well as developed data pre-processing methods. These methods are a revision from those presented in Vincenzi et al. (2022) where contamination is found to be less than 1.4 per cent for photometrically classified samples. We find that photometric quality is key for robust classification, and an improved sample can be expected from using high-quality Scene Modelling Photometry (Brout et al. 2019).

For future surveys such as LSST, photometric classification will be key to fully harness the power of these surveys. Photometric classification with host redshift information will enable using large, low-contamination, high-quality samples for measuring cosmological parameters. Potentially, MC BNN could provide useful information to filter transient samples in large surveys. Extensions to this work include photometric classification without redshift, which will assist in the allocation of follow-up resources for host galaxy redshift acquisition (such as Time-Domain Extragalactic Survey TiDES; Frohmaier et al. in preparation; Swann et al. 2019) and for other astrophysical studies.

ACKNOWLEDGEMENTS

AM thanks of the University of Queensland for hospitality during early work on this project; parts of this research were supported by the Australian Research Council Australian Laureate Fellowship FL180100168.

This paper has gone through internal review by the DES collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule Zürich (ETH), Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NFS's NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-

0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2).

This work was completed in part with Midway resources provided by the University of Chicago's Research Computing Center.

This work makes use of data acquired at the Anglo-Australian Telescope, under program A/2013B/012. We acknowledge the traditional owners of the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present.

MS is funded by the European Reearch Council (ERC) under the European Union's Horizon 2020 Research and Innovation program (grant agreement no 759194 - USNAC). LG acknowledges financial support from the Spanish Ministerio de Ciencia e Innovación (MCIN), the Agencia Estatal de Investigación (AEI) 10.13039/501100011033, and the European Social Fund (ESF) 'Investing in your future under the 2019 Ramón y Cajal program RYC2019-027683-I and the PID2020-115253GA-I00 HOSTFLOWS project, and from Centro Superior de Investigaciones Científicas (CSIC) under the PIE project 20215AT016. LK thanks the UKRI Future Leaders Fellowship for support through the grant MR/T01881X/1.

DATA AVAILABILITY

We provide in https://github.com/anaismoller/DES5YR_SNeIa_hostz: (i) the SNANA and/or Pippin configuration files to reproduce simulations in this paper, (ii) configuration files and scripts to retrain SNNclassifications models (SNN is an open source framework available in GitHub), and (iii) analysis code in python to reproduce plots and results. Sample classification probabilities are available in Zenodo https://doi.org/10.5281/zenodo.5904368.

REFERENCES

Abbott T. M. C. et al., 2018, ApJS, 239, 18
Angus C. R. et al., 2019, MNRAS, 487, 2215
Astier P. et al., 2006, A&A, 447, 31
Bernstein J. P. et al., 2012, ApJ, 753, 152
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Betoule M. et al., 2014, A&A, 568, A22
Brout D., Scolnic D., 2021, ApJ, 909, 26
Brout D. et al., 2019, ApJ, 874, 106
Butter A., Finke T., Keil F., Krämer M., Manconi S., 2022, JCAP, 4, 023
Caldeira J., Nord B., 2020, preprint (arXiv:2004.10710)
Campbell H. et al., 2013, ApJ, 763, 88
Carrasco Kind M., Brunner R. J., 2014, MNRAS, 442, 3380
Charnock T., Perreault-Levasseur L., Lanusse F., 2020, preprint (arXiv:2006.01490)
Childress M. J. et al., 2017, MNRAS, 472, 273
Coil A. L. et al., 2011, ApJ, 741, 8
Contreras C. et al., 2010, AJ, 139, 519
Dai M., Kuhlmann S., Wang Y., Kovacs E., 2018, MNRAS, 477, 4142
Dark Energy Survey, 2019, ApJ, 872, L30
de Jaeger T. et al., 2020, MNRAS, 495, 4860
DeGroot M. H., Fienberg S. E., 1983, J. R. Stat. Soc. Ser. D (The Statistician), 32, 12
Dietterich T. G., 2000, in Multiple Classifier Systems. Springer, Berlin, Heidelberg, p. 1
Flaugher B. et al., 2015, AJ, 150, 150
Foley R. J. et al., 2018, MNRAS, 475, 193
Fortunato M., Blundell C., Vinyals O., 2017, preprint (arXiv:1704.02798)
Frieman J. A. et al., 2008, AJ, 135, 338
Frohmaier C. et al., 2019, MNRAS, 486, 2308
Gal Y., Ghahramani Z., 2016, Proceedings of The 33rd International Conference on Machine Learning. PMLR, p. 1050
Goldstein D. A. et al., 2015, AJ, 150, 82
Grayling M. et al., 2021, MNRAS, 505, 3950
Gupta R. R. et al., 2016, AJ, 152, 154
Gutiérrez C. P. et al., 2020, MNRAS, 496, 95
Guy J. et al., 2007, A&A, 466, 11
Guy J. et al., 2010, A&A, 523, A7
Hicken M. et al., 2009, ApJ, 700, 331
Hicken M. et al., 2012, ApJS, 200, 12
Hinton S., Brout D., 2020, J. Open Source Softw., 5, 2122
Hlozek R. et al., 2012, ApJ, 752, 79
Hochreiter S., Schmidhuber J., 1997, Neural Comput., 9, 1735
Inserra C. et al., 2021, MNRAS, 504, 2535
Izmailov P., Vikram S., Hoffman M. D., Wilson A. G., 2021, preprint (arXiv:2104.14421)
Jha S. W., 2017, in Alsabti A. W., Murdin P., eds, Handbook of Supernovae. Springer International Publishing AG, p. 375
Jha S., Riess A. G., Kirshner R. P., 2007, ApJ, 659, 122
Jones D. O. et al., 2017, ApJ, 843, 6
Jones D. O. et al., 2018, ApJ, 857, 51
Kelsey L. et al., 2021, MNRAS, 501, 4861
Kessler R. et al., 2009, PASP, 121, 1028
Kessler R. et al., 2010, PASP, 122, 1415
Kessler R. et al., 2013, ApJ, 764, 48
Kessler R. et al., 2015, AJ, 150, 172
Kessler R. et al., 2019a, PASP, 131, 094501
Kessler R. et al., 2019b, MNRAS, 485, 1171
Kessler R., Scolnic D., 2017, ApJ, 836, 56
Kim E. J., Brunner R. J., Carrasco Kind M., 2015, MNRAS, 453, 507
Kunz M., Bassett B. A., Hlozek R. A., 2007, Phys. Rev. D, 75, 103508
Lakshminarayanan B., Pritzel A., Blundell C., 2017, Advances in Neural Information Processing Systems. Curran Associates, Inc., p. 30, preprint (arXiv:1612.01474)
Li W. et al., 2011, MNRAS, 412, 1441
Lidman C. et al., 2020, MNRAS, 496, 19
LSST Science Collaboration, 2009, preprint (arXiv:0912.0201)
Madau P., Weisz D. R., Conroy C., 2014, ApJ, 790, L17
Mandel K. S., Narayan G., Kirshner R. P., 2011, ApJ, 731, 120
Mandel K. S., Scolnic D. M., Shariff H., Foley R. J., Kirshner R. P., 2017, ApJ, 842, 93
Möller A., de Boissière T., 2019, MNRAS, 491, 4277
Möller A. et al., 2016, J. Cosmol. Astropart. Phys., 2016, 008
Möller A. et al., 2021, MNRAS, 501, 3272
Muthukrishna D., Narayan G., Mandel K. S., Biswas R., Hložek R., 2019, Publications of the Astronomical Society of the Pacific, 131, 1005
Nicolas N. et al., 2021, A&A, 649, A74
Pierel J. D. R. et al., 2018, PASP, 130, 114504
Popovic B., Brout D., Kessler R., Scolnic D., Lu L., 2021, ApJ, 913, 49
Pursiainen M. et al., 2018, MNRAS, 481, 894
Qu H., Sako M., Möller A., Doux C., 2021, AJ, 162, 67
Rest A. et al., 2014, ApJ, 795, 44
Sako M. et al., 2011, ApJ, 738, 162
Sako M. et al., 2018, PASP, 130, 064002
Scolnic D., Kessler R., 2016, ApJ, 822, L35
Scolnic D. et al., 2020, ApJ, 896, L13
Shivvers I. et al., 2017, PASP, 129, 054201
Smith M. et al., 2018, ApJ, 854, 37
Smith M. et al., 2020, AJ, 160, 267
Stritzinger M. D. et al., 2011, AJ, 142, 156
Swann E. et al., 2019, The Messenger, 175, 58
Villar V. A. et al., 2019, ApJ, 884, 83
Villar V. A. et al., 2020, ApJ, 905, 94

Vincenzi M., Sullivan M., Firth R. E., Gutiérrez C. P., Frohmaier C., Smith M., Angus C., Nichol R. C., 2019, MNRAS, 489, 5802
Vincenzi M. et al., 2020, MNRAS, 505, 2
Vincenzi M. et al., 2022, MNRAS, preprint ((arXiv:2111.10382)
Wilson A. G., Izmailov P., 2020, preprint (arXiv:2002.08791)
Wiseman P. et al., 2020a, MNRAS, 495, 4040
Wiseman P. et al., 2020b, MNRAS, 498, 2575
Wiseman P. et al., 2021, MNRAS, 506, 3330
Yuan F. et al., 2015, MNRAS, 452, 3047

APPENDIX A: UNCERTAINTIES AND FITTED PARAMETERS

In Section 5.2, we explored the interpretability of BNN uncertainties. We concluded that this interpretation was not straight forward from our results. Here, we extend this discussion by exploring possible correlations with other light-curve properties derived from a SALT2 fit in Fig. A1.

In general, we find that uncertainties tend to be larger for the data when compared with simulations. The uncertainties in BBB method varies more with the parameters.

We note that the classification uncertainties are large for red and high stretch SNe in the DES 5-yr sample. The median classification probability is also lower for these candidates. If the uncertainties are epistemic due to a smaller training set, then they would be large for the ends of the normal SNe Ia SALT2 parameter distributions since training sets have fewer such candidates. However, we do not find this behaviour. Another possible effect could be that bluer SNe Ia are more easily standardizable as previous literature suggests and thus their classification is more robust (Brout & Scolnic 2021; Kelsey et al. 2021). However, as this tendency is only observed in data and not simulations, no conclusion can be confidently drawn.

The peak magnitude in *i*-band behaviour in data agrees with that of the SNR of the light curve. Brighter candidates are classified with higher confidence than fainter ones. However, as in the previous Section we do not see such a behaviour in the simulation.

While the correlation between supernova properties and classification uncertainties are interesting to explore, they are difficult to interpret since multiple effects could be contributing to the uncertainties. Tests based on simple physical systems could provide hints towards further interpretability, such as recent work by Caldeira & Nord (2020).

APPENDIX B: DES 5-YR PHOTOMETRICALLY SELECTED SNE IA

A table with photometrically classified SNe Ia from all selection methods with their respective probabilities for a subsample of DES 5-yr data is provided at https://doi.org/10.5281/zenodo.5904368. Samples are selected using P larger than 0.5 for each method plus selection cuts.

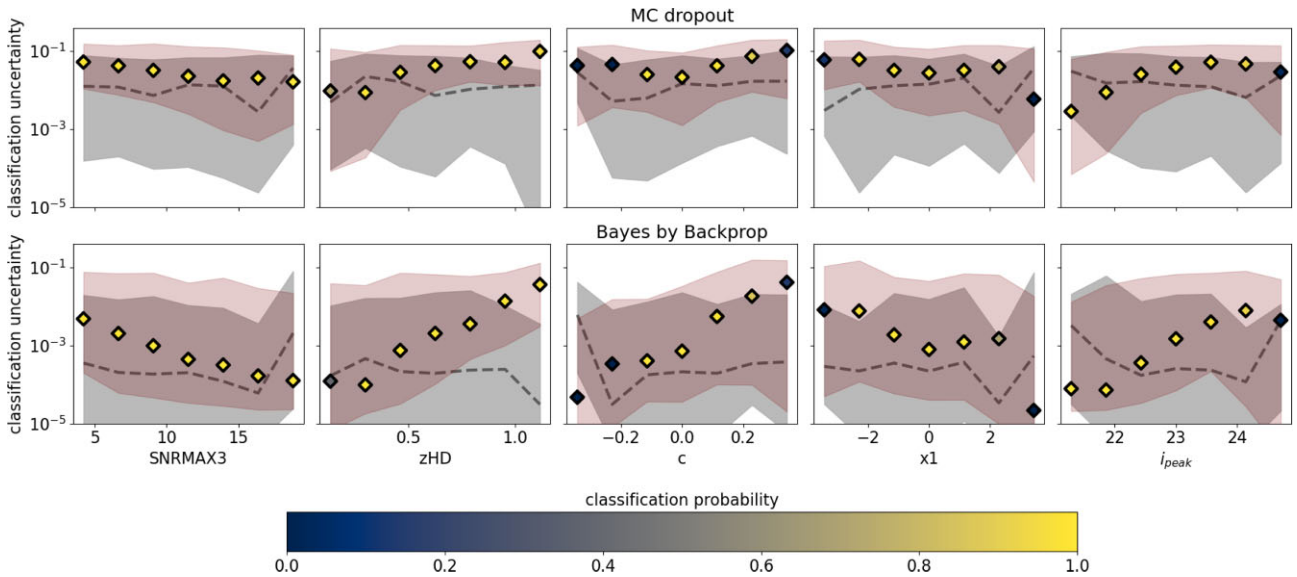

Figure A1. Distribution of classification uncertainties for DES 5-yr data (maroon) and simulation (grey) using the two BNNs, MC dropout and Bayes by Backprop. We show uncertainties as a function of SNR of the third brightest point in the light curve, redshift, colour, stretch and peak *i*-band magnitude. The median and 68 percentile are shown as a dashed line and filled coloured area. Data is shown as diamonds coloured by the median probability in that parameter bin and a maroon coloured area representing the 68 per cent percentile of the distribution.

Table B1. Example of SNNclassification probabilities for DES 5-yr candidates. A full list of classification probabilities for all DES 5-year candidates that are selected for any of the samples in this paper is found in https://doi.org/10.5281/zenodo.5904368. We show probabilities for baseline Recurrent Neural Network (RNN) in Section 4) and Bayesian Neural Networks methods MC dropout (MC) and Bayes by Backprop (BBB) in Section 5. For each method we provide classification probabilities rounded in two decimals for five different SNN initialization seeds, $s_0 = 0$, $s_1 = 55$, $s_2 = 100$, $s_3 = 1000$, $s_4 = 30469$, and the ensemble average probability of these five seeds *set*0.

IAUC	RNN						MC						BBB					
	s_0	s_1	s_2	s_3	s_4	*set*0	s_0	s_1	s_2	s_3	s_4	*set*0	s_0	s_1	s_2	s_3	s_4	*set*0
DES17S2gpk	0.99	0.95	0.99	0.99	1.00	0.98	1.00	0.96	0.98	0.98	0.99	0.98	0.97	1.00	0.92	0.99	0.98	0.97
DES14S2bck	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES14S2anv	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES15S2mji	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES13X3woy	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES14S2aoi	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES13C3xhy	0.73	0.99	1.00	0.32	0.00	0.61	0.82	0.99	0.54	0.94	0.60	0.78	0.46	0.82	0.85	0.98	0.63	0.75
DES15X3dyt	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES14E1gvc	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES13S2ead	0.99	1.00	1.00	1.00	0.98	0.99	1.00	0.98	1.00	1.00	0.99	0.99	0.99	1.00	0.98	0.98	0.98	0.99
DES16S1byw	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16E2bp	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES15E1nzd	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16X3enk	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16X3hy	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16X3hi	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES15C2mcu	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES13X1hxs	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES13X1bama	0.99	0.98	0.99	0.99	1.00	0.99	0.95	0.94	0.92	0.96	0.94	0.94	0.88	0.88	0.66	0.93	0.89	0.85
DES17C2acb	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16C3bab	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES17E2elx	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES14E2fyd	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16C3bq	0.97	0.99	0.79	0.31	1.00	0.81	0.95	0.83	0.91	0.94	1.00	0.93	0.78	0.95	0.69	0.94	0.87	0.85
DES15C3mes	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES15C3meu	1.00	0.99	1.00	1.00	0.98	0.99	1.00	0.97	0.90	0.89	0.99	0.95	1.00	0.98	0.89	1.00	0.98	0.97
DES16X3brw	0.79	0.02	0.35	0.02	0.25	0.29	0.99	0.94	0.96	0.97	0.71	0.92	0.91	0.95	0.64	0.86	0.73	0.82
DES14X1tbo	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00
DES15X3dyv	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16E2dcg	0.71	0.01	0.50	0.33	0.56	0.42	0.95	0.98	0.94	0.95	0.93	0.95	0.88	0.86	0.75	0.69	0.96	0.83
DES16E2dch	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES15C2iuv	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES13X2gnl	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16E1byy	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table B1 – *continued*

IAUC	RNN						MC						BBB					
	s_0	s_1	s_2	s_3	s_4	*set*0	s_0	s_1	s_2	s_3	s_4	*set*0	s_0	s_1	s_2	s_3	s_4	*set*0
DES15E2mhj	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES14E2cmo	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES14E2hhu	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES13E1aftw	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES14C2ocp	1.00	1.00	1.00	0.98	1.00	0.99	0.99	1.00	0.99	1.00	0.99	0.99	1.00	0.99	0.83	0.99	0.95	0.95
DES13X2jdk	1.00	0.97	0.99	0.99	0.99	0.99	0.99	0.97	0.99	0.98	0.96	0.98	0.98	0.98	0.96	0.95	0.98	0.97
DES14E2clm	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES17C3blq	0.01	1.00	0.96	0.32	0.99	0.66	0.94	1.00	0.79	0.35	0.29	0.67	0.93	0.75	0.96	0.96	0.18	0.76
DES16E2blm	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES15X3auw	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES15X2mey	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES15C3odz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES14C3oce	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.73	0.99	0.89	0.99	0.98	0.92
DES15X2mfa	0.99	0.99	0.99	0.89	0.99	0.97	0.99	0.97	0.97	0.94	0.95	0.96	0.93	0.98	0.94	0.93	0.96	0.95
DES16S2buz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16C3fhz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES17C2emh	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES17E1bmf	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES17C3ivv	0.87	1.00	0.63	0.96	1.00	0.89	0.81	0.99	0.98	0.92	0.90	0.92	0.99	0.99	0.99	0.98	1.00	0.99
DES17E1blu	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97	1.00	0.98	0.98	0.98
DES14S2frj	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES17E2bmb	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES17C1ify	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DES16X1drk	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

[1]*Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Victoria 3122, Australia*
[2]*LPC, CNRS/IN2P3, Université Clermont Auvergne, F-63000 Clermont-Ferrand, France*
[3]*CNRS/IN2P3, Université de Lyon, Université Claude Bernard Lyon 1, IP2I Lyon, F-69622 Villeurbanne, France*
[4]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[5]*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*
[6]*Department of Physics, Duke University Durham, NC 27708, USA*
[7]*The Research School of Astronomy and Astrophysics, Australian National University, ACT 2601, Australia*
[8]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), 28040 Madrid, Spain*
[9]*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
[10]*INAF-Osservatorio Astronomico di Trieste via Tiepolo 11, I-34143 Trieste, Italy*
[11]*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
[12]*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK*
[13]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain*
[14]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain*
[15]*Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA*
[16]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[17]*Sydney Institute for Astronomy, School of Physics, The University of Sydney, A28, NSW 2006, Australia*
[18]*Centre for Gravitational Astrophysics, College of Science, The Australian National University, ACT 2601, Australia*
[19]*Department of Physics, University of Surrey, Guildford, Surrey GU2 7XH, UK*
[20]*Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile*
[21]*Laboratório Interinstitucional de e-Astronomia – LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ 20921-400, Brazil*
[22]*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*
[23]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[24]*Institut d'Astrophysique de Paris, Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, F-75014 Paris, France*
[25]*Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr 1, D-81679 Munich, Germany*
[26]*Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[27]*Kavli Institute for Particle Astrophysics & Cosmology, Stanford University, P. O. Box 2450, Stanford, CA 94305, USA*
[28]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[29]*Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain*
[30]*Dpto. Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain*
[31]*Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*
[32]*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA*
[33]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra, Barcelona, Spain*
[34]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain*
[35]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08E-193 Barcelona, Spain*
[36]*Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[37]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK*
[38]*Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy*
[39]*INAF-Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*
[40]*Institute for Fundamental Physics of the Universe, Via Beirut 2, I-34014 Trieste, Italy*
[41]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ 20921-400, Brazil*
[42]*Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India*
[43]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
[44]*Institute of Theoretical Astrophysics, University of Oslo. P.O. Box 1029 Blindern, NO-0315 Oslo, Norway*
[45]*Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, E-28049 Madrid, Spain*
[46]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[47]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[48]*Excellence Cluster Origins, Boltzmannstr 2, D-85748 Garching, Germany*
[49]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
[50]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[51]*Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia*
[52]*Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA*
[53]*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA*
[54]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[55]*Physics Department, University of Wisconsin–Madison, 2320 Chamberlin Hall, 1150 University Avenue Madison, WI 53706-1390, USA*
[56]*Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA*
[57]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[58]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[59]*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton BN1 9QH, UK*
[60]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[61]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany*
[62]*Fakultät für Physik, Ludwig-Maximilians Universität München, Universitäts-Sternwarte, Scheinerstr 1, D-81679 München, Germany*

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