Classification of Variable Stars Light Curves Using Long Short Term Memory Network

Lunwei Zhang

2024-03-15

https://www.frontiersin.org/articles/10.3389/fspas.2021.718139/full

Outline

Part I Introduction

Part II Data and Method

Part III Results

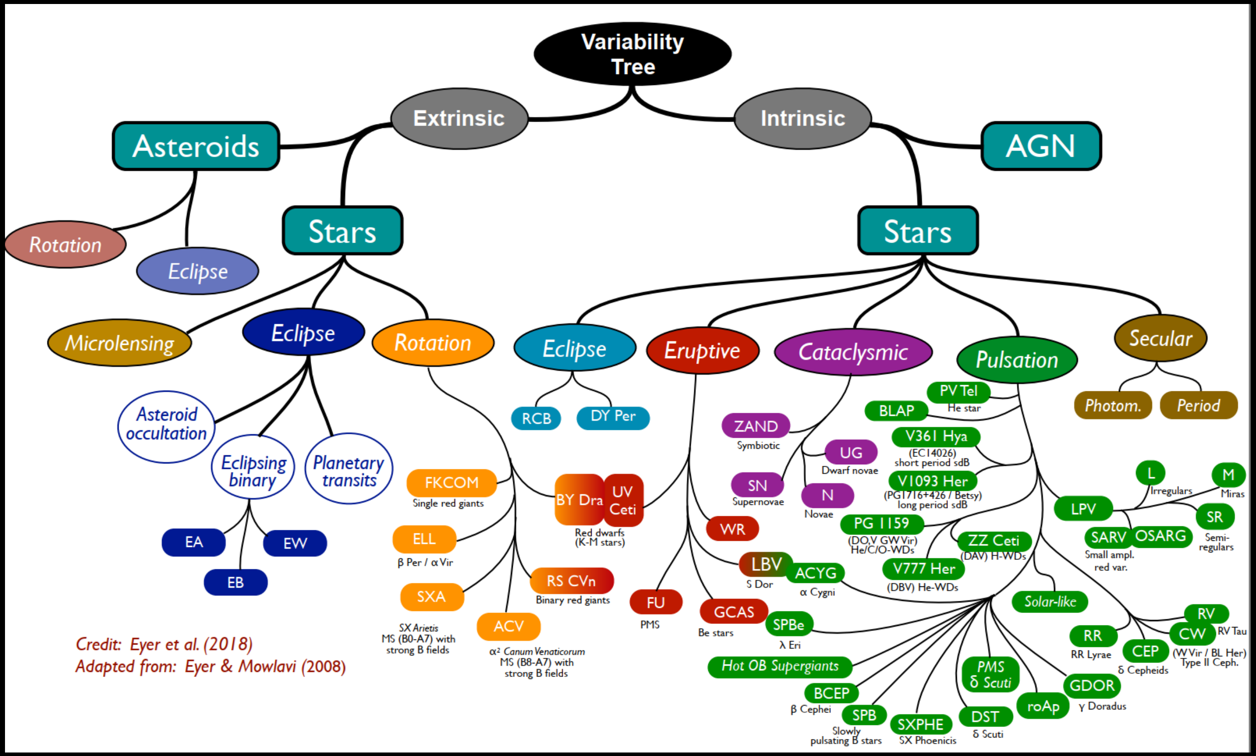
Part IV Summary and future work

Part I Introduction

Introduction

- Distance indicators, estimating distances to galaxies within and beyond the Local Group and measuring the Hubble constant
- Helps us to understand the evolution and physics of stars themselves(Stellar variability)
- Tracing new structures or better study the Galactic plane, the spiral arms, and the solar neighborhood
- Studying chemical composition of different galactic regions
- Studying planetary formation through premain-sequence stars

……



Time-domain surveys

- OGLE(Udalski et al., 1993; Soszyński., 2015, Soszyński., 2016, Soszyński., 2018),
- ASAS(Pojmanski, 2002),
- CRTS; (Drake et al., 2009;Djorgovski et al., 2011),
- ZTF(Bellm et al., 2019),
- Vera C. Rubin Observatory (LSST Science Collaboration et al., 2009)
- ……

Introduction-previous work

Traditional ML

- 28 features(Fourier analysis) Debosscher et al. (2007) ML classifiers
- statistical parameters Dubath et al. (2011) Random Forest
- combine the periodic features with the non-periodic features Richards et al. (2011)
- FATS Nun et al. (2015)
- UPSILoN Kim and Bailer-Jones (2016) Random Forest
- Support Vector Machines (SVM),Neural Nets (NN), Random Forests (RF) Pashchenko et al. (2018)

Deep Learning

- The raw-light curves to generate dm-dt maps Mahabal et al. (2017) CNN
- Unsupervised and effective feature extraction (Naul et al. (2018) RNN)
- 1D CNN (Aguirre et al. (2019))
- Sequence of images Carrasco-Davis et al. (2019) recurrent convolutional neural network (RCNN)
- Probabilistic classification model (Zorich et al. (2020))

Feature based (Directly or hierachical)
Raw-data
Phase-fold image
Image+features (multimodal)

Introduction-Motivations

- Classifying these sources based on their light curves helps us in understanding the responsible mechanisms behind the variability and provides insight into their interior structure and formation;

- Development of automated methods for classifying variable star’s light curves has seen an upward trend in recent years and has formed the core of many latest studies;

- Time-series data might be sparse (and therefore not good enough to estimate the period) and can contain gaps in the observations;

- Recent studies focus on employing the raw time-series data and take advantages from the improved deep-learning (DL) frameworks.

Part II Data and Method

2.1 Data

Only a few of classes which have enough number of distinct light curves (**~500 or more**) selected

Class	Representation	Number
OGLE dataset		
Classical Cepheids	CEPH	2698
δ Scuti	DSCT	464
Eclipsing Binaries	ECLP	50,000
Long Period Variables	LPV	22,371
RR Lyrae	RRL	28,473
CRTS dataset		
Contact Binaries	EW	30,745
Long Period Variables	LPV	511
Detached Binaries	EA	4683
RR Lyrae type 1	RRab	2431
RR Lyrae type 2	RRc	28,473
RR Lyrae type 3	RRd	502
Rotating Variables	RSCVn	1522

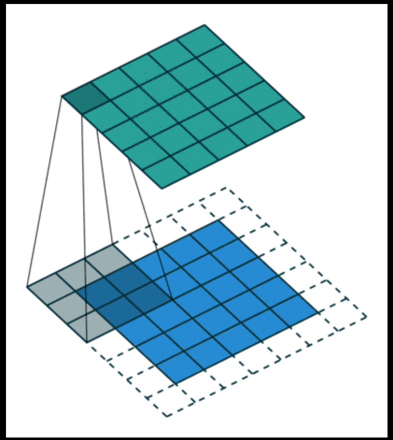
104,006

68,867

60% training set,20% validation set 20% test set

Using the two datasets separately for training and testing the 2D CNN and 1D CNN-LSTM models!!!

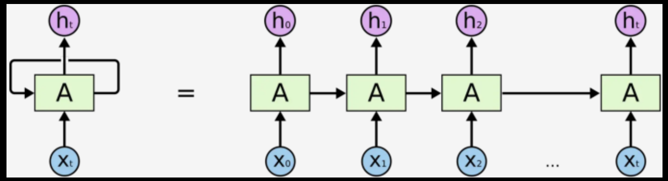
2.2 Method



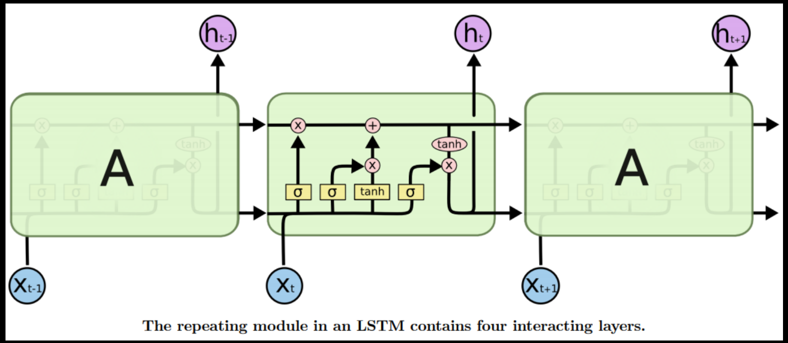
Convolution

CNN

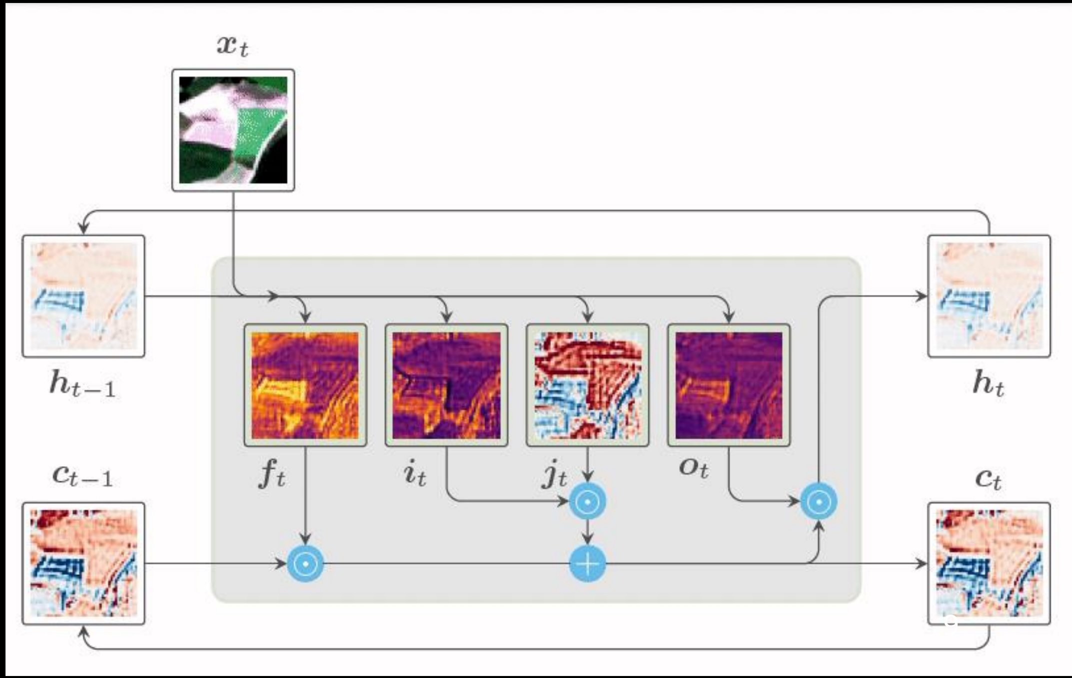
Credit from the Internet



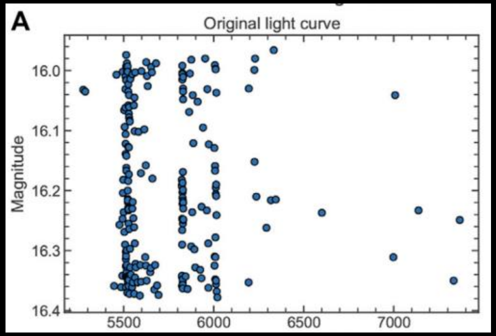
RNN



LSTM



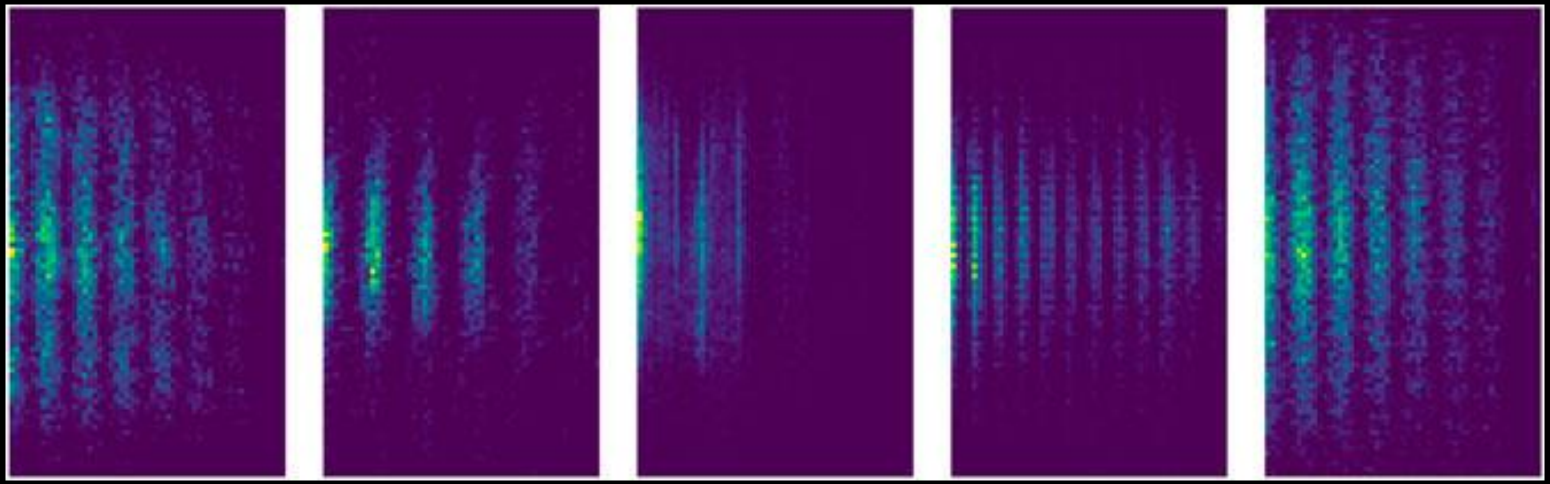
2.2 Method-Pre-processing



Raw light curve data

Freedman-Diaconis rule

53 along x-axis and 90 along y-axis((R, G, and B))



Cepheids δ-Scuti EB LPV RR Lyrae

Freedman-Diaconis (Freedman and Diaconis, 1981)

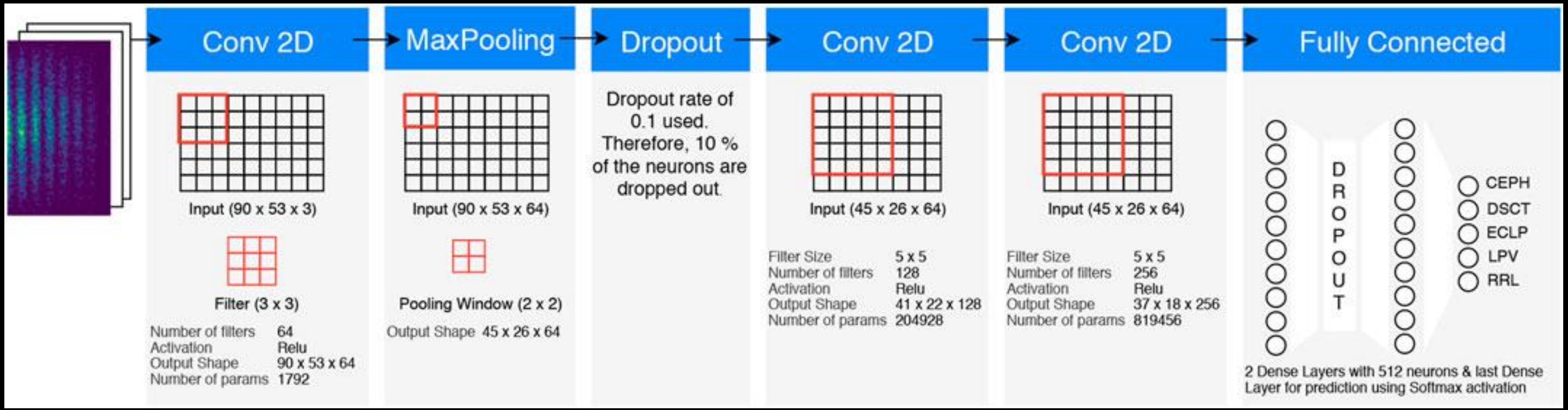
Representative bi-dimensional histograms generated from the OGLE light curves. The blue and yellow color represent the pixels with the least and the most number of points respectively.

2.2 Method

1/2 CNN

OGLE

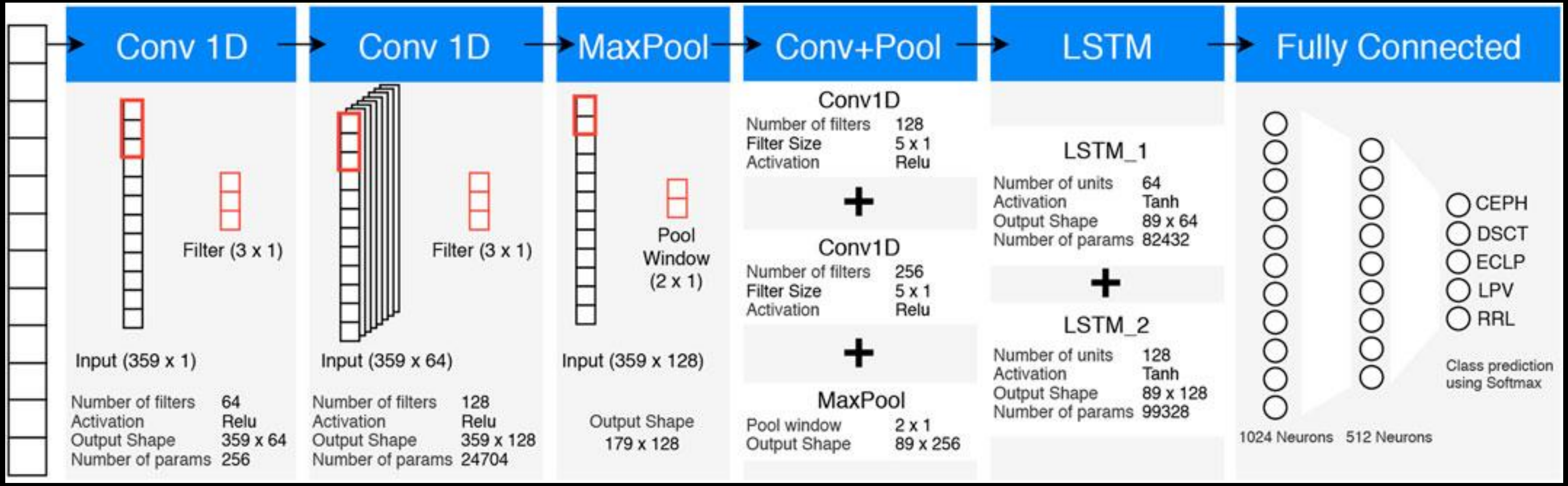
$90 \times 53 \times 3$



2/2 1DCNN LSTM

OGLE 359

CRTS 546



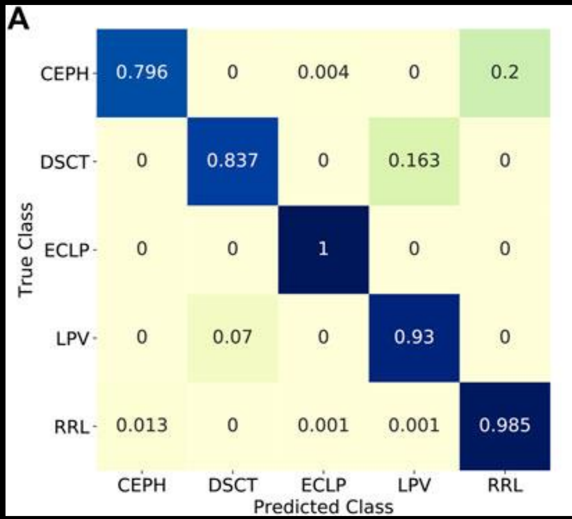
Part III Results and Discussion

3.1 Results

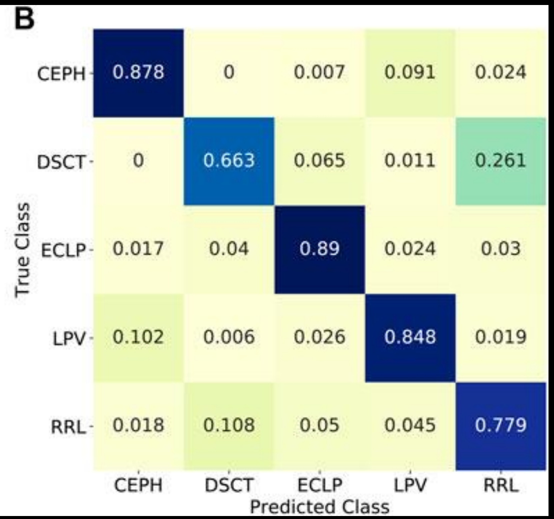
2D CNN on bi-dimensional histograms prepared from the OGLE and the CRTS survey light curves

Dataset	Accuracy (%)	Precision	Recall	F1 score
OGLE	97.5	0.81	0.91	0.85
CRTS	74.5	0.56	0.52	0.54

2D CNN perform very well on the OGLE dataset but suboptimalon the CRTS dataset



2D CNN(OGLE)



1D CNN-LSTM (OGLE)

2D CNN perform better 1D CNN-LSTM on the OGLE

1D CNN-LSTM on the OGLE and the CRTS

Dataset	Accuracy (%)	Precision	Recall	F1 score
OGLE	85.0	0.64	0.81	0.71
CRTS	66.6	0.46	0.53	0.49

1D CNN-LSTM perform well on the OGLE dataset but suboptimalon the CRTS dataset

Comparison of classification results for three variability classes from the CRTS dataset using 2D CNN and 1D CNN-LSTM models.

True class	Predicted class	Classification models	
		2D CNN	1D CNN-LSTM
RRab	RRab	29%	55%
	EW	51%	22%
RRc	RRc	42%	45%
	EW	50%	41%
LPV	LPV	69%	80%
	RRab	13%	6%

The superiority of the 1D LSTM-CNN model over 2D CNN in distinguishing very similar looking light curves

Part IV Conclusion and future work

4.1 Summay

- Present two approaches for classifying variable stars using Deep Learning techniques;
- For both 2D CNN and 1D CNN-LSTM, the classification performance on the CRTS dataset is suboptimal;
- The performance of 1D CNN-LSTM model is not at par with the 2D CNN approach;
- The degraded performance on the CRTS dataset as compared to the OGLE dataset is a common difficulty faced by both the models;
- 1D CNN-LSTM model has the potential to perform the task of classifying variable star light curves without any preprocessing.

4.2 Future work

- Investigating the other binning strategies along with the ones proposed in Mahabal et al. (2017)
- Performing detailed comparisons with the bidimensional histogram, 1D LSTM-CNN approaches and other prevailing classification techniques
- Exploring the capability of the hyperparameter optimized model in classifying light curves from different surveys and examining their performance in case of the sparse light curves
- Using a combination of two parallel CNNs, a 1D CNN for the light curves and another 2D CNN for the science (or difference) images

Thank you for your attention!
Q&A

Research Status

- Traditionally, **based on the similarity of their light curves and colors to known variable prototypes** (Gaia Collaboration et al. 2019).
- **Time-series analysis: periodogram** to differentiate between periodic and aperiodic variables.
- The results of the **periodogram-based analysis** are often taken, together with measures of **light curve morphology** and **other characteristics of the source (e.g., color)** and used as inputs into a classifier (e.g. Debosscher et al. 2007; Richards et al. 2011; Dubath et al. 2011; Richards et al. 2012; Masci et al.2014; Jayasinghe et al. 2019a,b; Eyer et al. 2019).
- **Non-parametric variability measures** (Kinemuchi et al. 2006; Palaversa et al.2013; Drake et al. 2013, 2014a, 2017; Torrealba et al. 2015; Hillenbrand &Findeisen 2015; Findeisen et al. 2015)

Classical Methods: carefully selected features of the light curves, such as statistical metrics (like mean, standard deviation, kurtosis, skewness; see e.g., Nun et al. 2015), Fourier decomposition (Kim & Bailer-Jones 2016), or color information (Miller et al. 2015). Classifiers can be trained on manually designed (Pashchenko et al. 2018; Hosenie et al. 2019) or computer-selected features (Becker et al. 2020; Johnston et al. 2020) using known types of variable stars

Research Status

F. Förster et al. 2020

Table B3. Light curve based ML classifiers that include only persistent variable objects (more than 2 classes) before 2017. Class abbreviations are defined in Tables B7 to B11

Reference	Data source	#classes	classes
Kim & Bailer-Jones (2016)	MACHO, LINEAR, ASAS	19	DSCT, RRL(ab, c, d, e), Ceph(F, O1, other, II), E(C, SD, D), LPV(MAGBC, MAGBO, OSARGAGB, OSARGRGB, SRAGBC, SRAGBO), NV
Mackenzie et al. (2016)	OGLE	6	Ceph(CL, II), RRL, E, DSCT, LPV
	MACHO	8	NV, QSO, BeS, Ceph, RRL, E, ML, LPV
Pichara et al. (2016)	MACHO	8	BeS, Ceph, E, LPV, ML, NV, QSO, RRL
	EROS	11	E, RRL, Ceph(F, O1, DM, II), LPV(OSARGRGBO, SRAGBO, SRAGBC, MAGBC, MAGBO)
Nun et al. (2016)	MACHO	8	NV, QSO, BeS, Ceph, RRL, E, ML, LPV
Bass & Borne (2016)	*Kepler*	14	ACT, BCep, Ceph, DSCT, E, ELL, GDor, ROT, RRL(ab, c), RVTau, SPB, SR, MISC/NV
Faraway et al. (2016)			
Kügler et al. (2015)	OGLE	3	Ceph, E, RRL
	ASAS	7	Mira, RRLab, E(C, D, SD), DSCT, CephF
Kim et al. (2014)	EROS-2	26	DSCT, RRL(ab, c, d, e), Ceph(F, O1, Other), CephII E(C, SD, D, SD+D, Other), BeS, QSO, NV LPV(MAGB(C, O), OSARGAGB(C, O), OSARGRGB(C, O), SRAGB(C, O))
Pichara & Protopapas (2013)	SAGE, 2MASS, UBVI, MACHO	7	NV, QSO, BeS, Ceph, RRL, E, LPV
Richards et al. (2012)	ASAS	28	DSCT, SXPh, RRL(ab, c, d), Ceph(CL, MM, II), Mira, SR, LPVW(A, B), RVTau, BCep, RSG, BPer, BLyr, WUMa, ChemPec, ELL, RSCvn, HAeBe, CTTau, WLTTau, RCB, LBV, BeS
Debosscher et al. (2009)	CoRoT	29	sdBV, DSCT, LBoo, SXPh, roAp, GDor, RR(ab, c, d), Ceph(CL, DM, II), RVTau, Mira, SR, PVSG, BCep, SPB, E, ChemPec, ELL, FUOri, HAeBe, TTau, LBV, WR, XB, BeS, LAPV
Debosscher et al. (2007)	OGLE	35	DAV, DBV, sdBV, GWVir, DSCT, LBoo, SXPh, roAp, GDor, RRL(ab, c, d), Ceph(Cl, DM, II), PVSG, Mira, SR, RVTau, BCep, SPB, E(C, SD, D), ChemPec, ELL, FUOri, HAeBe, TTau, LBV, SLR, WR, XB, CV, BeS

Table B2. Light curve based ML classifiers that include only persistent variable classes (more than 2 classes) between 2017 and 2019. Class abbreviations are defined in Tables B7 to B11

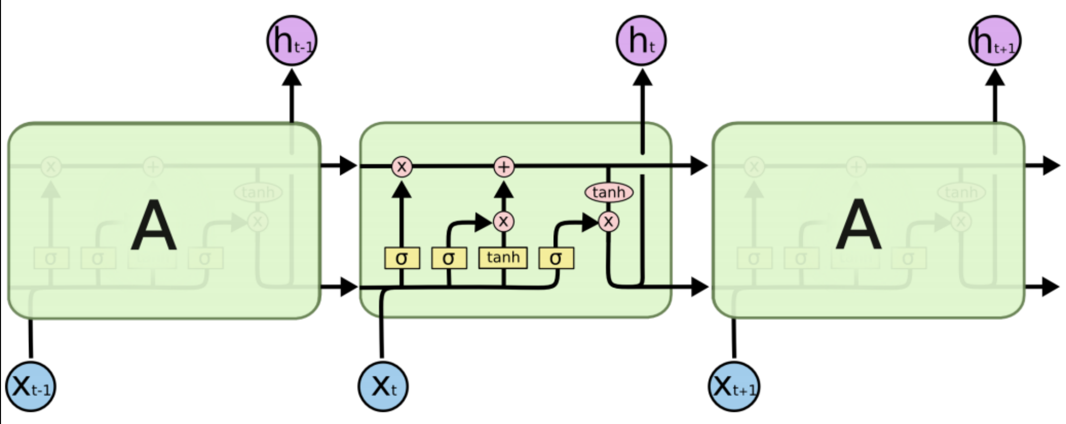
Reference	Data source	#classes	classes
Rimoldini et al. (2019)	*Gaia* DR2	18	E, CV, RSCvn, BLAP, Mira+SR, DSCT+SXPh, RRL(ab, c, d, Ad), CephCl, ACEP, CephII, Low amp.:DSCT+GDOR, ELL, OSARG, FL+ROT, Other
Tsang & Schultz (2019)	ASAS-SN	8	DSCT, RRL(ab, cd), Ceph, E, ROT, Mira, SR
Jayasinghe et al. (2019)	ASAS-SN	10	Ceph, DSCT, E(EW,EA—EB,EB), RRL(ab,c), M, SR, Irregular
Hosenie et al. (2019)	CSDR2	12	RRL(ab, c, d), Blazhko, E(C+SD,D), ROT, LPV, DSCT, Ceph(II,A)
Johnston et al. (2019)	UCR	3	RRL, Ceph, E
	LINEAR	5	RRL(ab, c), DSCT, E(C,SD)
Aguirre et al. (2019)	OGLE+VVV +CoRoT	9	Ceph(F, 01), RRL(ab, c), E(C, SD+D), Mira, SR, OSARG
Castro et al. (2018)	MACHO	8	NV, QSO, BeS, Ceph, RRL, E, ML, LPV
	OGLE	6	Ceph, CephII, RRL, E, DSCT, LPV
Naul et al. (2018)	ASAS	5	RRLab, Ceph, SR, BPer, WUMa
	LINEAR	5	DSCT, RRL(ab, c), BPer, WUMa
	MACHO	8	Ceph(F, O1), LPVW, RRL(ab, c, e, GB)
Valenzuela & Pichara (2018)	OGLE	8	Ceph(CL, II, A), RRL, LPV, DPV, DSCT, E
	MACHO	11	RRL(ab, c, e, GB), Ceph(F, O1), LPVW(A, B, C, D), E
Mahabal et al. (2017)	CSDR2	7	E(C, SD), RRL(ab, c, d), RSCVn, LPV
Benavente et al. (2017)	EROS, MACHO, HiTS	5	Ceph, E, QSO, RRL, LPV
Zinn et al. (2017)	OGLE	8	Mira, QSO, SR, OSARG, Ceph(F, O1), RRL(ab+d, c+e)

Table B5. Observational data sources used for ML classification.

Abbreviation	Long name	Reference
ZTF	Zwicky Transient Facility	Bellm et al. (2019)
HSC–SSP	Hyper Suprime-Cam Subaru Strategic Program	Aihara et al. (2018)
UCR	University of California Riverside Time Series Classification Archive	Dau et al. (2018)
OSC	Open Supernova Catalog	Guillochon et al. (2017)
ASAS-SN	All-Sky Automated Survey for Supernovae	Kochanek et al. (2017)
CSDR2	The Catalina Surveys Data Release 2	Drake et al. (2017)
HiTS	High cadence Transient Survey	Förster et al. (2016)
PS1-MDS	Pan-STARRS-1 Medium Deep Survey	Huber et al. (2011)
LINEAR	Lincoln Near-Earth Asteroid Research Survey	Sesar et al. (2011)
UBVI	UBVI photometry of six open cluster candidates	Piatti et al. (2011)
VVV	Vista Variables in the Via Lactea	Minniti et al. (2010)
OGLE	The Optical Gravitational Lensing Experiment	Udalski et al. (2008)
2MASS	The Two Micron All Sky Survey	Skrutskie et al. (2006)
SAGE	Spitzer Survey of the Large Magellanic Cloud: Surveying the Agents of a Galaxy's Evolution	Meixner et al. (2006)
CoRoT	Convection, Rotation, and planetary Transits	Baglin et al. (2006)
SDSS	The Sloan Digital Sky Survey	York et al. (2000)
MACHO	Massive Compact Halo Objects survey	Alcock et al. (2000)
EROS	Expérience pour la Recherche d'Objets Sombres	Palanque-Delabrouille et al. (1998)
ASAS	All Sky Automated Survey	Pojmanski (1997)

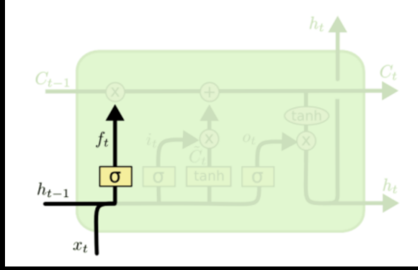
F. Förster et al. 2020

LSTM



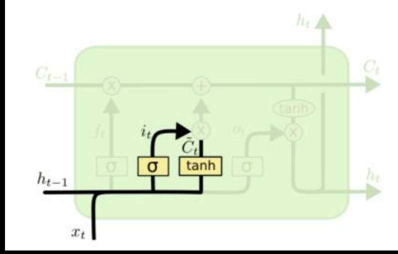
The repeating module in an LSTM contains four interacting layers.

Forget Gate



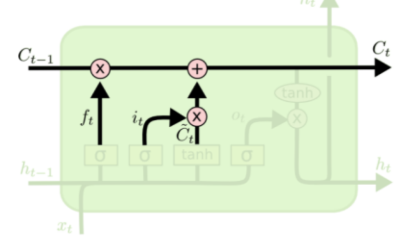
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right)$$

Input Gate



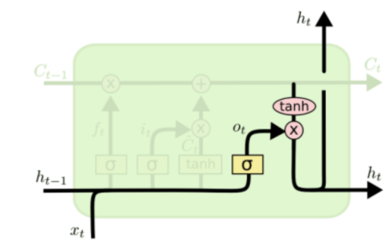
$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Memory Cell



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Output Gate



$$o_t = \sigma\left(W_o \, [h_{t-1}, x_t] + b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$