

Semi-Supervised domain adaptation for multi-band Supernovae classification

J. Saavedra-Bastidas¹, D. Moreno-Cartagena², M. Pérez-Carrasco³ and G. Cabrera-Vives²

¹Astronomy Department, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Concepción, Chile

²Department of Computer Science, Universidad de Concepción, Concepción, Chile

³Data Science Unit, Universidad de Concepción, Concepción, Chile

Keywords: (transients:) supernovae, methods: statistical, methods: data analysis, techniques: photometric

Corresponding author: jsaavedra2018@udec.cl

We introduce a novel approach that extends the Minimax Entropy (MME) algorithm for the multiband photometric light curve classification of supernovae. We assess the performance of our classifier based on the number of objects provided during training, the number of observations during prediction, and the multiband capabilities of the top candidates. Our results demonstrate the effectiveness of our method in the early classification of Zwicky Transient Facility (ZTF) transients.

In this work, we utilize two datasets: transients obtained through the ALerCE broker with forced photometry from ZTF and synthetic light curves (SLC). Both datasets include fluxes for type Ia, II, Ibc, and super-luminous supernovae in the *g* band (~ 464 nm) and *r* band (~ 658 nm), along with temporal information regarding the first alert per object.

We defined two sets of experiments: a *fine-tuning* strategy that involves taking a deep learning model previously trained with SLC data and further training it using a smaller labeled set from the ZTF sample, and a *domain adaptation* experiment through Minimax Entropy (Saito et al., 2019). This experiment utilized the full simulated dataset, a small number of labeled items from the ZTF sample, and an unlabeled dataset comprising the remaining ZTF objects. We optimized the model parameters via a two-term loss function:

$$\mathcal{L} = H(y_s, \hat{y}_s) + \lambda H(\hat{y}_u) \quad (1)$$

Here, the first term represents the cross-entropy loss between the true and predicted values, and the second term denotes the entropy of the predicted labels for the unlabeled dataset.

The results for various sequential models (Muthukrishna et al., 2019; Bai et al., 2018; Astorga et al., 2023) on the 15th day after the trigger, sorted by ZTF objects delivered during training, are presented in Table 1. The Temporal Convolutional Network (TCN) proves to be the most suitable for classification, outperforming the Recursive

Table 1. Metric score, 15th day after the first alert.

Shots	Train	F1-Score(%)		
		RNN	TCN	TLC
1	FT	49±8	52±2	39±7
	MME	49±7	54±7	27±4
5	FT	50±5	51±4	44±5
	MME	49±7	55±3	41±5
10	FT	49±6	55±3	46±4
	MME	54±6	55±5	36±7
15	FT	52±5	53±3	44±7
	MME	55±2	51±7	37±5
20	FT	44±4	51±7	40±2
	MME	50±7	50±6	38±2

Neural Network (RNN) and the Transformer (TLC). The MME algorithm is the most effective, requiring only five ZTF objects per class during training, surpassing fine-tuning (FT), which needs a minimum of ten objects. When evaluating the top TCN classifiers on the SLC dataset on the 15th day after the trigger, MME achieves $52\pm3\%$, outperforming FT ($51\pm3\%$) and remaining competitive with the fully-trained TCN on SLC data ($55\pm1\%$).

Acknowledgments: We extend our gratitude to Dr. Daniel Muthukrishna and the ALerCE team for providing access to the data used in this study. We also acknowledge the support of the ANID BASAL project FB21003.

References

- Astorga, N., et al. 2023, Research Square, doi: [10.21203/rs.3.rs-2395110/v1](https://doi.org/10.21203/rs.3.rs-2395110/v1)
- Bai, S., Kolter, J. Z., & Koltun, V. 2018, arXiv e-prints. <https://arxiv.org/abs/1803.01271>
- Muthukrishna, D., Narayan, G., Mandel, K. S., Biswas, R., & Hložek, R. 2019, PASP, 131, 118002, doi: [10.1088/1538](https://doi.org/10.1088/1538)

-[3873/ab1609](#)

Saito, K., Kim, D., Sclaroff, S., Darrell, T., & Saenko, K.

2019, in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 8050–8058