2° Forum della Ricerca Sperimentale e Tecnologica

Deep Learning for Photometric Metallicity Estimation of RR Lyrae Stars Using Gaia Data

Lorenzo Monti, Tatiana Muraveva, Gisella Clementini, Alessia Garofalo
INAF - OAS

1 INTRODUCTION

Background:
Astronomy has entered the era of Big Data, driven by missions such as **ESA's Gaia telescope**, which has produced an extensive catalogue of over 270,000 RR Lyrae stars in its **Gaia Data Release 3** (DR3). These stars are invaluable for tracing the oldest stellar populations and measuring distances in the Milky Way and nearby galaxies. Gaia's DR3 provides time-series photometry, radial velocity data, and other essential parameters for a large and homogeneous sample of RR Lyrae stars (***Clementini et al., 2023***), making them central to studies of the Milky Way's structure and evolution.

Problem Statement:
Metallicity (*[Fe/H]*) is a crucial factor in understanding stellar evolution and galactic chemical composition. Traditionally, metallicity has been derived from Fourier parameters of RR Lyrae light curves. Traditional empirical methods to predict RR Lyrae stars' metallicity often introduce biases due to data heterogeneity and noise. A deep learning approach offers a more accurate prediction using light curves.

Objective:
Use deep learning to estimate the metallicity of fundamental mode (RRab) and first-overtone (RRc) RR Lyrae stars from their Gaia DR3 G-band light curves. By harnessing the power of deep learning, we aim to push the boundaries of astronomical research and unlock new insights into the chemical and structural evolution of our galaxy.

2 METHODS

Dataset: taken from **Gaia DR3** catalogue, we selected a set of **6002 RRab stars** and a set of **6613 RRc stars** based on: *err[Fe/H] < 0.4dex; ϕ_{31}error < 0.10; peak-to-peak amplitude < 1.4 mag; number of epochs > 50;*.

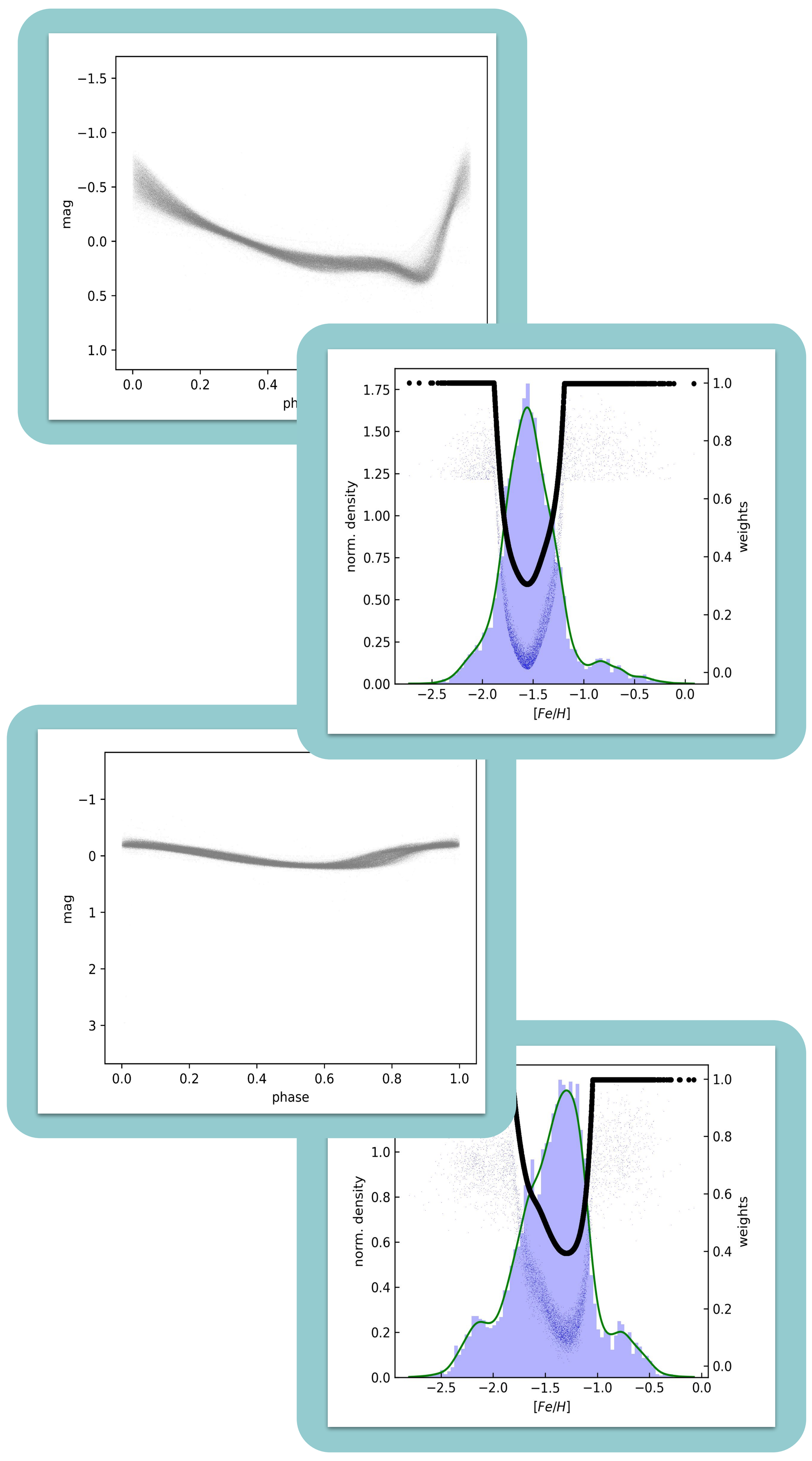
Predictive Modeling: for the predictive modeling of the [Fe/H] from the light curves, we use the following two dimensional sequences as input variables:

$$X^{<t>} = \begin{cases} m^{<t>} - <m> \\ Ph * P \end{cases} \quad t = 1, \ldots, N_{ep}$$

where $m^{\langle t \rangle}$is the magnitude of the light curve, *<m>* is the mean magnitude, *Ph* is the phase and *P* the period. N_{ep} is the number of epochs.

Preprocessing phase: preprocessing involved **phase folding** and **smoothing spline** technique to minimize fluctuations, noise, outliers and obtain the same number of points for each light curve (*264*). Moreover, we applied **sample weights** for metallicity distribution. We computed *Gaussian kernel density* estimates of the [Fe/H] distributions. Finally, we evaluated them for each object in the datasets and assigned a density weight w_d to each data point by taking the inverse of the estimated normalized density.

Deep Learning Models: nine models were tested, including **CNNs**, **RNNs**, and **hybrid architectures**. Stratified **K-fold cross-validation** (5 folds) was performed to evaluate model performance. For both types of RR Lyrae stars, the **GRU model** (3 layers) showed the best performance with *L1* and *L2 regularization*, *dropout layers*, and *Adam optimizer*. For more details read the paper (***Monti et al, 2024***).

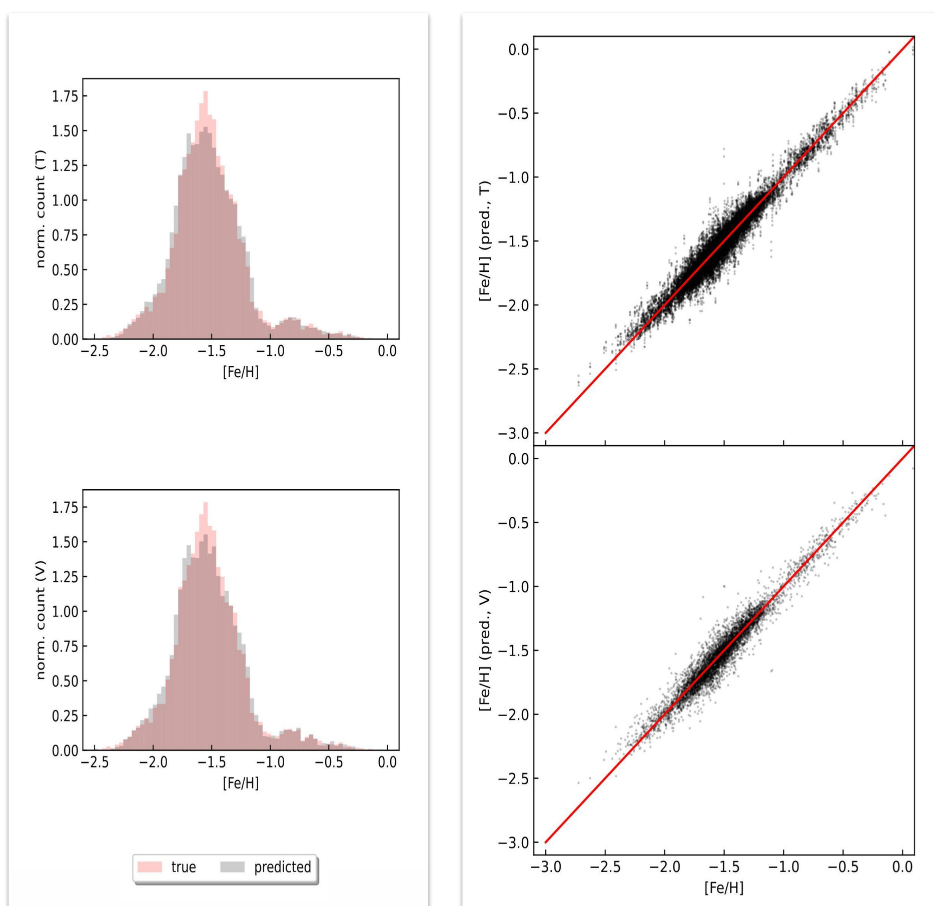


Phase-folded light curves and metallicity distributions for RR Lyrae stars. On top: fundamental mode (RRab) and on bottom first-overtone (RRc) RR Lyrae stars.

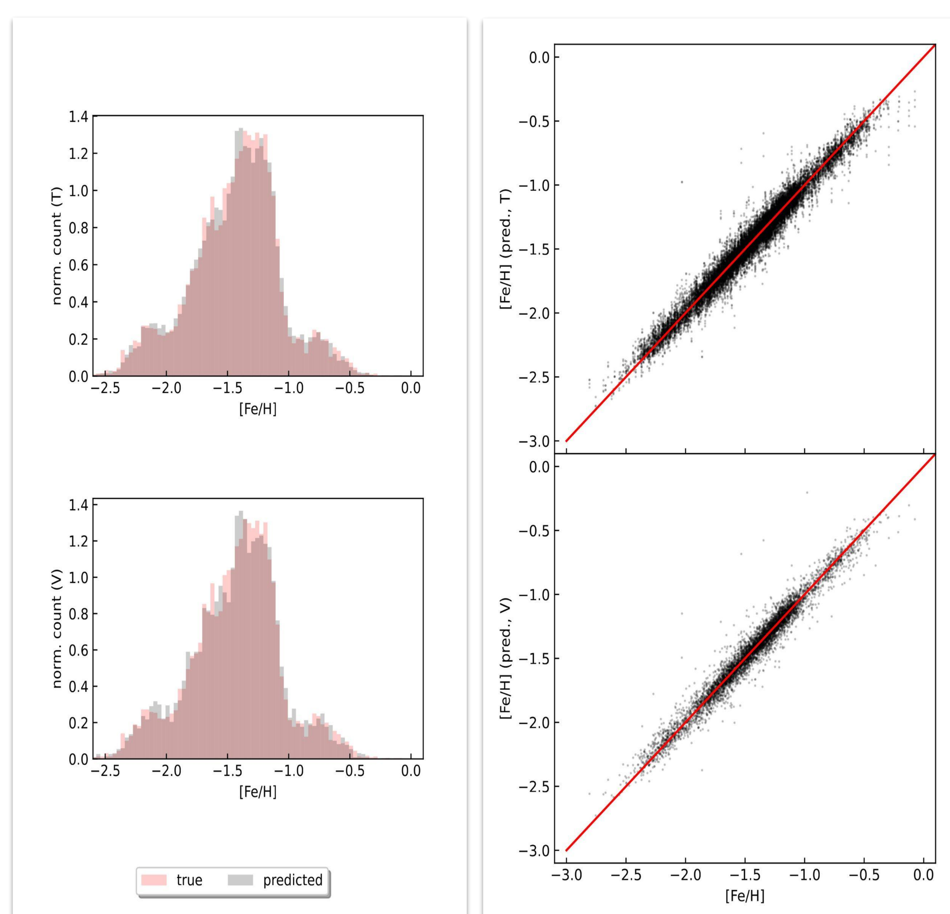
3 RESULTS

The metrics suggest minimal generalization error and strong predictive capabilities. The table compares model performance metrics for **RRab** and **RRc** stars, showing a good balance between bias and variance for both models. Preprocessing improved model performance, with **GRU-based models** performing best. Moreover, the figures present histograms and scatter plots of **ground-truth vs. predicted [Fe/H] values** for both models, demonstrating high accuracy and effective generalization.

GRU Model	Model Performance RRab stars	Model Performance RRc stars
R^2 (Validation)	0.9401	0.9625
RMSE (Validation)	0.0765	0.0720
MAE (Validation)	0.0565	0.0505



RRab stars



RRc stars

4 DISCUSSION

Advances: The GRU-based model achieves significant improvement over traditional methods in predicting RR Lyrae stars' metallicity. Deep learning effectively captures complex relationships between photometric light curve features and metallicity.

Next steps: (1)new architectures like **Transformers** applied to time series. (ii) approaches used for Natural Language Processing such as **seq2seq** and applying them to time series.

5 CONCLUSION

Deep learning, particularly **GRU-based models**, offers precise metallicity predictions for RR Lyrae stars both for fundamental mode and first-overtone RR-Lyrae. The integration of such models in astrophysics can significantly enhance our understanding of stellar populations.

REFERENCES

Clementini et al., 2023. Clementini, G.; Ripepi, V.; Garofalo, A.; Molinaro, R.; Muraveva, T.; Leccia, S.; Rimoldini, L.; Holl, B.; de Fombelle, G.J.; Sartoretti, P.; et al. Gaia Data Release 3-Specific processing and validation of all-sky RR Lyrae and Cepheid stars: The RR Lyrae sample. Astron. Astrophys. **2023**, 674, A18

Monti et al., 2024. Lorenzo Monti, Tatiana Muraveva, Gisella Clementini, Alessia Garofalo. "Leveraging Deep Learning for Time-Series Extrinsic Regression in Predicting the Photometric Metallicity of Fundamental-Mode RR Lyrae Stars ." MDPI Sensors 24.16 (2024) p. 5203.

REPOSITORY

CONTACT

Lorenzo Monti
lorenzo.monti@inaf.it