From Fermi-LAT Unassociated Sources to Pulsar Discoveries: A Hierarchical Deep Learning 1D-CNN Approach for Spectral Classification

C. Pozo González, R López-Coto, J Mendez-Gallego, E. De Ona

Instituto de Astrofísica de Andalucía - CSIC

Cristian Pozo González

ML4Astro / 01-09-2026

Credits: Gemini

VHEGA

IAA INSTITUTO DE ASTROFÍSICA DE ANDALUCÍA

EXCELENCIA SEVERO OCHOA

Content

- Instrument Fermi-LAT
- Our scope
- Network Architecture
- Result
  - Spatial Crosscheck
  - ATNF crossmatch
  - FAST detections
- Conclusion

Paper accepted for publication in A&A

29/07/2026

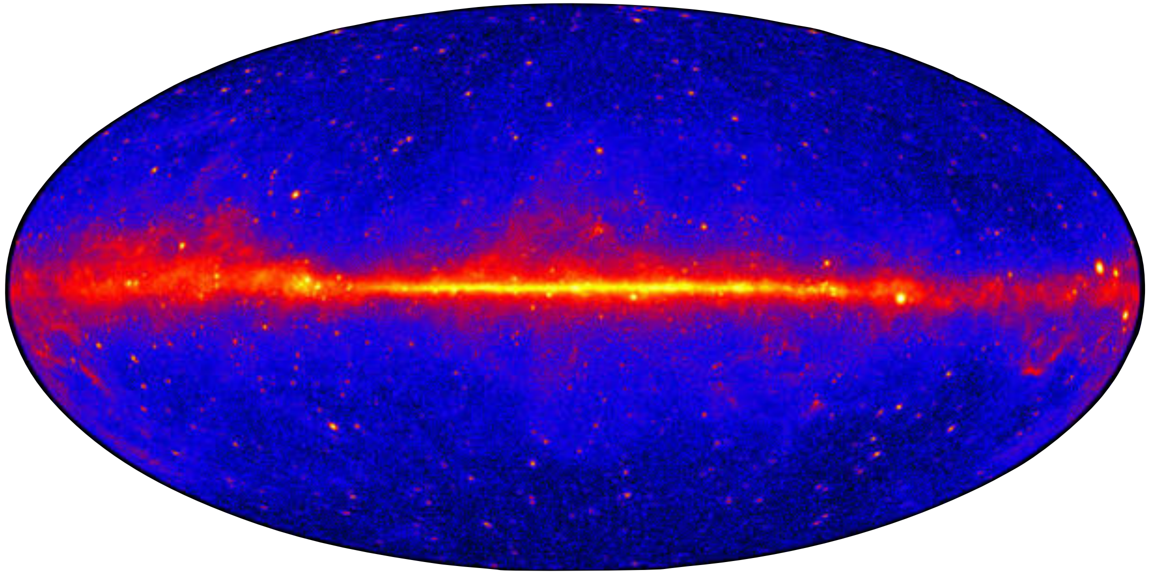
Deep Learning-Based Classification and Analysis of Pulsar Candidates in Fermi-LAT Unassociated Sources

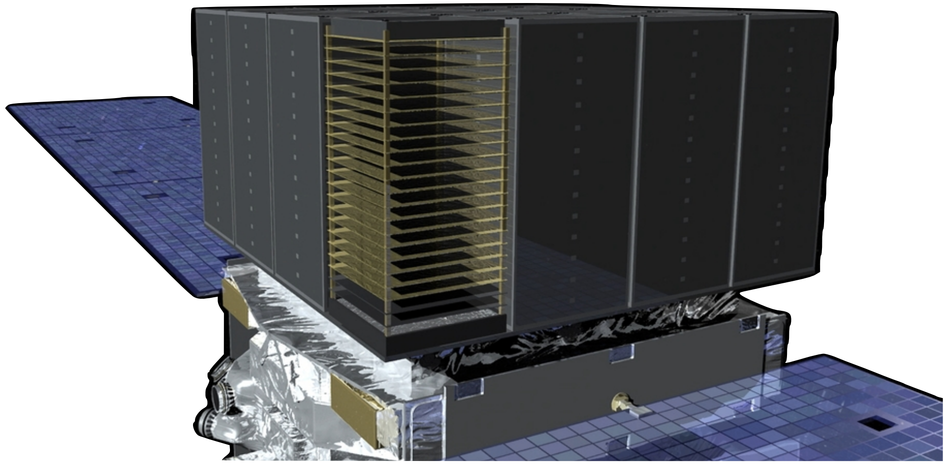
C. Pozo González, R. López-Coto, J. Méndez-Gallego, E. de Oña Wilhelm

The Large Area Telescope (LAT) has revolutionized our understanding of the high-energy sky, yet approximately one-third of the sources in the Fourth Fermi-LAT Source Catalog (4FGL) remain unassociated. Conventional machine learning, such as Decision Trees, often treat spectral features as independent tabular entries, neglecting the sequential topological information inherent in the Spectral Energy Distribution (SED). We aim to classify unassociated Fermi-LAT sources by exploiting the intrinsic shape of their spectra and variability features, avoiding the use of galactic coordinates as training features. Our primary objective is to generate a high-confidence list of PSRs candidates, further distinguishing between Young Pulsars (YPs) and Millisecond Pulsars (MSPs). We developed a hierarchical deep learning framework based on a 1D Convolutional Neural Network (1D-CNN), named TabularResCNN. This architecture treats the spectral data from the 4FGL catalog, allowing the model spectral shape. The classification is performed in two stages: first discriminating between AGNs and PSRs, and subsequently categorizing PSRs into YPs and MSPs. We implemented a cost-sensitive learning strategy to handle class imbalance and utilized Grad-CAM techniques to ensure the physical interpretability of the model's decisions. Applying this framework to 2563 unassociated sources, we identified 1136 AGNs and 202 high-confidence PSR candidates (166 YPs, 36 MSPs), increasing the pulsar population by more than 60%. They exhibit strong astrophysical consistency: YPs are confined to the Galactic plane, MSPs show a broader vertical distribution, and AGNs are isotropic. Furthermore, we identified 5 out of 5 PSRs recently confirmed by FAST. The proposed 1D-CNN framework isolates PSRs candidates based on intrinsic spectral and temporal properties, minimizing spatial bias.

Comments: Accepted in Astronomy & Astrophysics (A&A)
Subjects: **High Energy Astrophysical Phenomena (astro-ph.HE)**
Cite as: arXiv:2607.28723 **[astro-ph.HE]**
(or arXiv:2607.28723v1 **[astro-ph.HE]** for this version)
https://doi.org/10.48550/arXiv.2607.28723

Fermi-LAT instrument





From 2008 is continue doing survey of the Gamma ray (50 MeV - 1 TeV) sky

DRL4 (14 years) provide more that 5000 sources

More than 2000 unassociated sources

Our scope

Class	Description	Sources	Percentage
Training & Validation Classes			
BLL	BL Lac object	1,468	20.40%
FSRQ	Flat Spectrum Radio Quasar	776	10.79%
BCU	Blazar of Uncertain type	1,622	22.54%
AGNs Total	**(BLL + FSRQ + BCU)**	**3,866**	**53.73%**
YPs	Young Pulsar	141	1.96%
MSPs	Millisecond Pulsar	179	2.49%
PSRs Total	**(PSR + MSP)**	**320**	**4.45%**
Target for Prediction			
UNK	Unassociated Sources	2,423	33.68%

Classify the unassociated sources with a hicherical classification:

- ❖ First into two mayor class AGNs and PSRs
- ❖ Second isolate YPS and MSPs

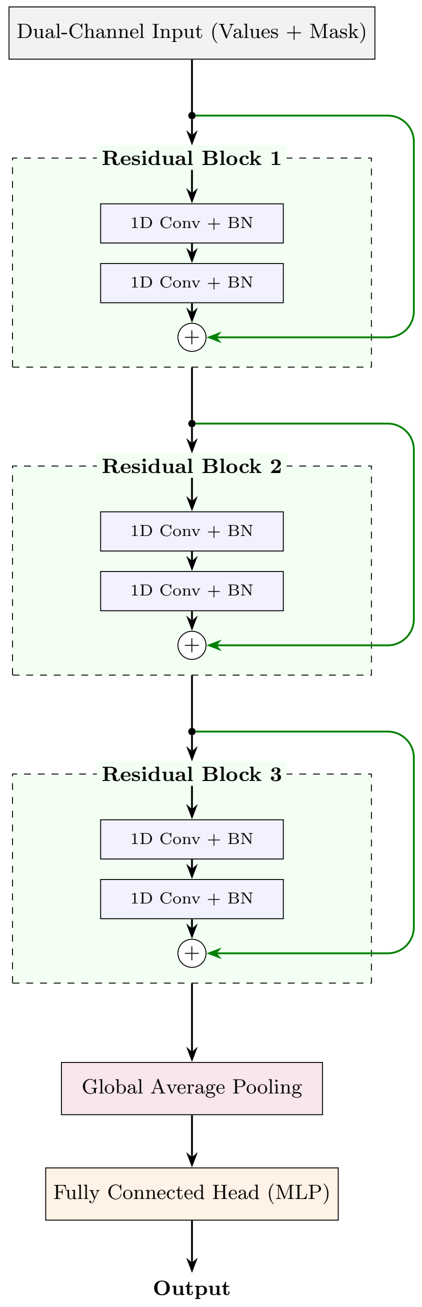
The method: TabularResCNN

- **1D Spectral Topology**: Treats the Spectral Energy Distribution (SED) as a sequential 1D signal.

- **Dual-Channel Input**: Combines normalized feature values with a binary sensitivity mask. Handles missing data and learns instrumental limits without artificial data manipulation.

- **Residual Architecture**: Uses 1D convolutional layers with skip connections.

- **Spatial Bias Elimination**: Explicitly excludes galactic coordinates from training features.



Paper and code repository in prep.

Result: Metrics

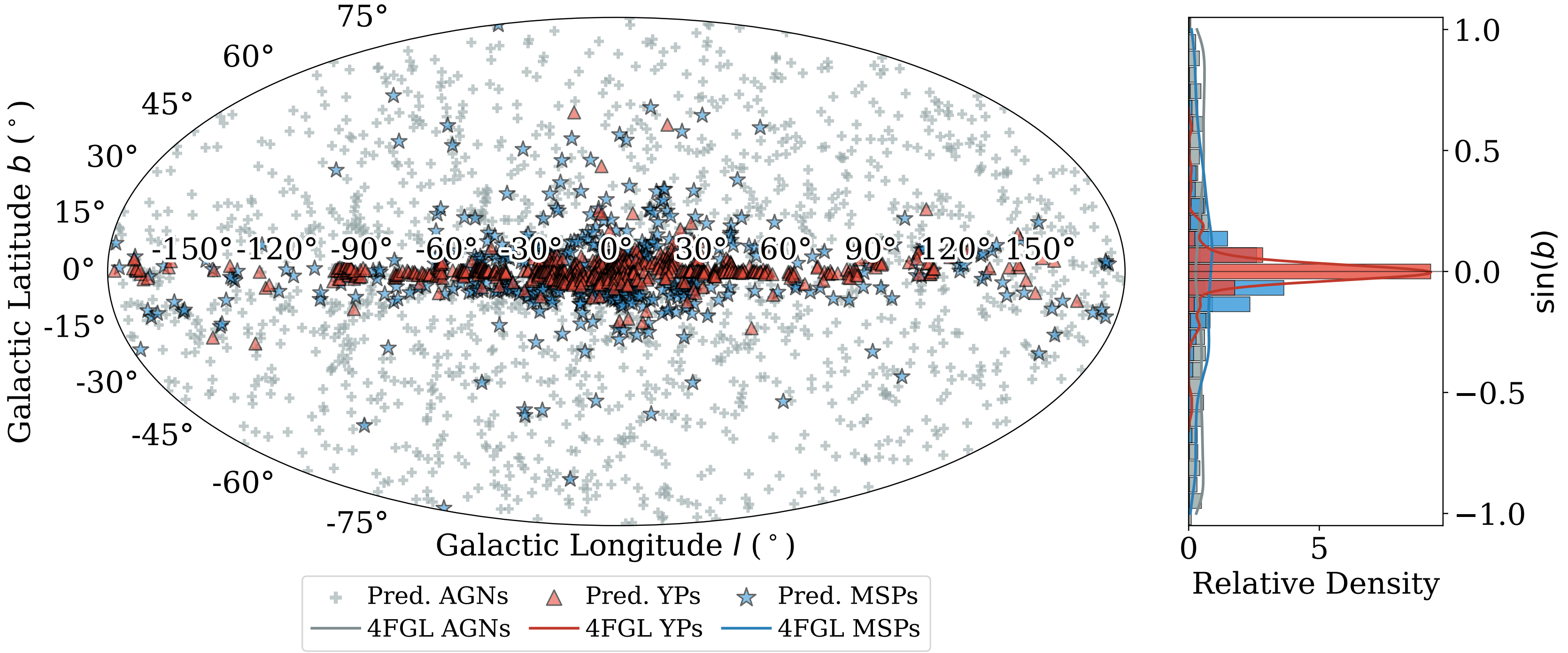
❖ **Stage 1 (AGNs vs. PSRs):** 96.7% overall accuracy Near-perfect AGN retrieval and high precision (95%) for pulsars.

Class	Precision	Recall	F1-Score	Support
AGNs	0.9853	0.9937	0.9895	473
PSRs	0.9500	0.8906	0.9194	64
Macro average	**0.9677**	**0.9421**	**0.9544**	**537**

❖ **Stage 2 (YPs vs. MSPs):** 81.4% accuracy Performance reflects the intrinsic spectral overlap between populations.

Class	Precision	Recall	F1-Score	Support
MSPs	0.8824	0.8108	0.8451	37
PSRs	0.7200	0.8182	0.7660	22
Macro average	**0.8012**	**0.8145**	**0.8055**	**59**

Result: Spatial Crosscheck



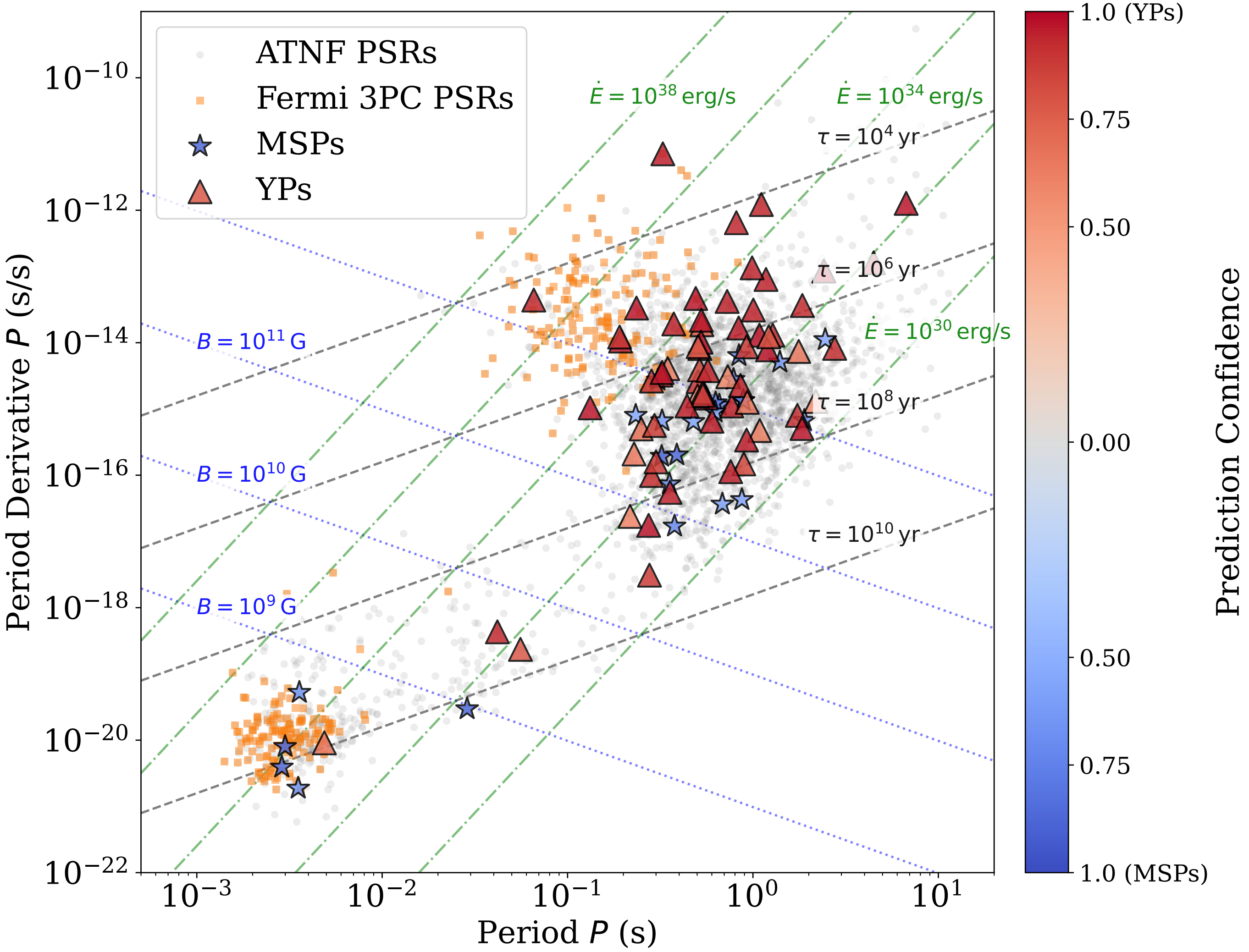
- ❖ Young Pulsars (YPs): Tightly confined to the Galactic plane
- ❖ Millisecond Pulsars (MSPs): Broader vertical distribution due to kinematic heating
- ❖ AGNs: Fully isotropic across the high-latitude sky.

Result: ATNF crossmatch

❖We performed a spatial crossmatch in of the PSRs candidate with ATNF pulsars.

❖ Our classification reproduce the MSPs and YPs timing features.

❖ We don't use any timing information during the training



Result: FAST detections

Source Name	ℓ (deg)	b (deg)	Class	$P_{\rm PSRs}$	$P_{\rm MSP}$	σ	$\mathrm{Flux}_{1\,\mathrm{GeV}}$ (10^{-11} ph cm^{-2} s^{-1})
4FGL J0237.8+5238	138.83	-6.92	MSP	0.98	0.85	26.9	129.5
4FGL J0533.6+5945	152.26	14.17	MSP	0.92	0.84	17.6	70.0
4FGL J1730.4−0359	19.74	16.00	MSP	0.98	0.86	21.6	142.5
4FGL J1827.5+1141	40.75	10.55	MSP	0.93	0.94	18.9	94.9
4FGL J1904.7−0708	28.07	-6.20	PSR	0.85	0.31	24.9	184.7

- ❖ **Predictive Power:** The candidate catalog was cross-referenced with recent independent blind-search pulsar discoveries by the FAST telescope.

- ❖ **Survey Prioritization:** This perfect agreement highlights the catalog's reliability as an optimized, bias-minimized roadmap for allocating observational time in future radio surveys (e.g., FAST, MeerKAT, SKAO).

Conclusion

Paper accepted for publication in A&A
29/07/2026

❖**Bias-Free Framework:** TabularResCNN successfully isolates pulsar candidates using intrinsic spectral and temporal properties, eliminating spatial coordinate bias.

❖ **Population Expansion:** Processed 2,563 unassociated sources, identifying 1,136 AGNs and 202 high-confidence PSR candidates (166 YPs, 36 MSPs) , increasing the pulsar census by >60%.

❖**Empirical Validation:** Demonstrated 100% accuracy (5/5) on recent blind-search discoveries by the FAST telescope prior to radio confirmation.

❖**Observational Roadmap:** Provides a highly refined, cost-sensitive target list optimized for maximum discovery yields in future surveys with FAST, MeerKAT, and SKAO.