Pulsar Detection with Deep Learning

Manideep Pendyala

November 7, 2023

Abstract

The rapid growth of data in astronomy necessitates automated approaches for pulsar candidate detection. This thesis employs deep learning techniques, specifically convolutional neural networks (CNNs), to identify and classify pulsar candidates. The models are trained on manually labeled data collected by the Giant Metrewave Radio Telescope (GMRT). The thesis aims to create a robust and accurate pulsar candidate detection system, addressing the challenge of handling large-scale pulsar survey data.

1 Understanding Pulsars

Pulsars are rapidly rotating neutron stars with intense magnetic fields ($\approx 10^{15}$G) that emit electromagnetic radiation. They are mostly observed in wavelengths outside the visible spectrum, primarily in radio wavelengths, and display a distinct pulsating pattern due to their regular rotations. This regularity in their emissions makes them a fascinating object of study in the field of astronomy and astrophysics.

1.1 Characteristics of Pulsars

- **Strong Magnetic Fields:** Pulsars have magnetic fields between one billion and one quadrillion times stronger than Earth's, leading to unique phenomena.
- **Radio Wave Emission:** Pulsars emit tightly focused beams of radio waves from their magnetic poles.
- **Rapid Rotation:** Pulsars rotate rapidly, akin to lighthouses, and their beams need to align with Earth for us to detect their pulses.
- **Precision Clocks:** Pulsars serve as highly precise cosmic clocks due to their stable and predictable pulse intervals, despite potential variations caused by various factors such as spin slowdown, companion stars, and external acceleration.

1.2 Significance of Pulsars

- **Precision Cosmic Clocks**: Pulsars serve as incredibly accurate cosmic timekeepers, rivaling Earth's best atomic clocks. Their predictability enables precise measurements in astrophysics.

- **Probing the Interstellar Medium**: Observing pulsars helps determine the density and distribution of interstellar matter, offering a unique window into cosmic environments impossible to replicate on Earth.
- **Space Navigation**: Pulsars, with their precise pulses, act as celestial beacons and may serve as a cosmic GPS for space navigation, providing reliable reference points for spacecraft, especially in deep space where traditional Earth-based systems are unreachable.
- **Testing theories of General relativity**: We can evaluate the validity of general relativity by observing how the orbits of celestial objects (particularly binary stars) gradually deteriorate as they emit gravitational waves and dissipate energy.

1.3 Pulsar Detection and Identification

Finding pulsars begins with the collection of radio wave data using powerful radio telescopes. However, due to the inherent weakness of individual pulses compared to background noise, direct detection of pulsars is challenging.

To overcome these challenges, astronomers utilize a combination of methods. One crucial technique is the Fourier transform, which converts the data from the time domain to the frequency domain. This transformation helps identify periodic signals within the data.

Folding is another essential step where astronomers take advantage of the known or suspected pulsar period to combine multiple weak pulses into a more robust,integrated pulse. This process significantly enhances the signal-to-noise ratio, making it possible to detect otherwise faint pulsars.

Specialized software suits like PRESTO assist in automating the Fourier transform and conducting preliminary searches for pulsar signals. However, human verification remains indispensable in distinguishing genuine pulsar signals from spurious candidates.

The process results in the generation of diagnostic plots (an example can be found Figure 1), which provide visual representations of the folded data. Astronomers scrutinize these plots to identify pulsar candidates. However, this is a challenging task, particularly in the context of the vast amount of data collected in modern pulsar observations Zhu et al. (2014).

2 Deep Learning in Astrophysics

Deep learning is a subset of machine learning, which is inspired by the structure and function of the human brain, particularly neural networks. These algorithms attempt to simulate the behavior of the human brain to process data and make decisions but at a much larger and more complex scale.

Astronomers have adopted deep learning techniques to automate the process of screening radio signals from pulsars, aiming to significantly improve the accuracy and efficiency of their observations. Deep learning models, particularly convolutional neural networks, offer a promising solution to address scalability issues when identifying and classifying pulsar

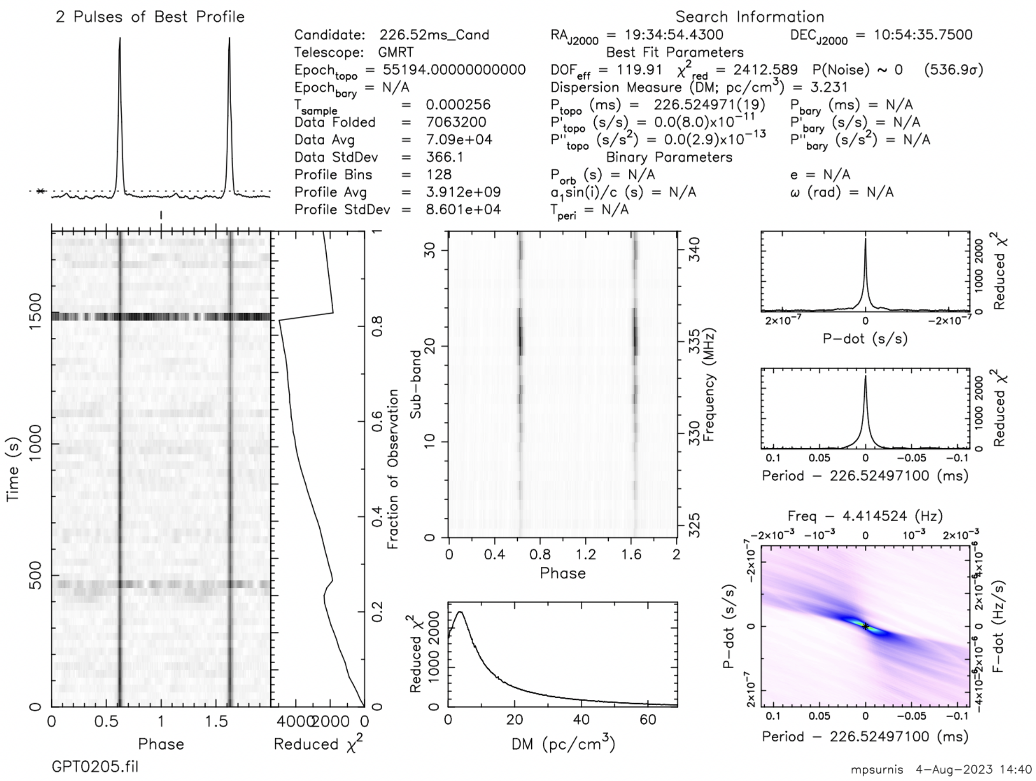

Figure 1: An example of diagnostic plot

candidates.

2.1 Deep Learning Framework Overview

The basis of our model uses a combination of 1D and 2D data from a pulsar candidate plot in Figure 1 to feed it through a logistic regression as shown in Figure 2, on which we will build further.

Input Data: Phase vs Frequency and Phase vs Time are taken as images. These images provide visual information related to the candidate. Summed pulse profile and DM (Dispersion Measure) vs SNR (Siganl to Noise ratio) graphs are in the form of 1D arrays of data, which are numerical measurements of time series data associated with the pulsar candidate. All of these correspond to the same candidate and will help our model gain a better understanding by providing diverse features.

1. **Image Processing:** - We take each of the two images and process them using a Convolutional Neural Network (CNN) and Support Vector Machines (SVMs). Each CNN and SVM will extract relevant features from its respective image.

2. **1D Array Processing:** - Similarly, We process each of the two 1D arrays using an Artificial Neural Network (ANN) designed for numerical data and SVMs.

3. **Combining Outputs:** After processing the images through the CNN and the 1D arrays through the ANN and SVM, you will have a total of eight different outputs.

To make a final classification decision, we use a Logistic Regression (LR) model to combine the outputs of the CNN, SVM, ANN, and SVM for both the images and 1D arrays.

Output: The output of the logistic regression model will be a classification decision, typically binary (e.g., pulsar or non-pulsar).

This approach allows you to leverage both deep learning (CNN for images, ANN for 1D arrays) and traditional machine learning (SVM for feature classification and logistic regression for combining outputs) to make a final classification decision based on multiple data sources. It's a multi-modal classification system that takes advantage of the unique strengths of each type of model to make an informed decision about the pulsar candidate.

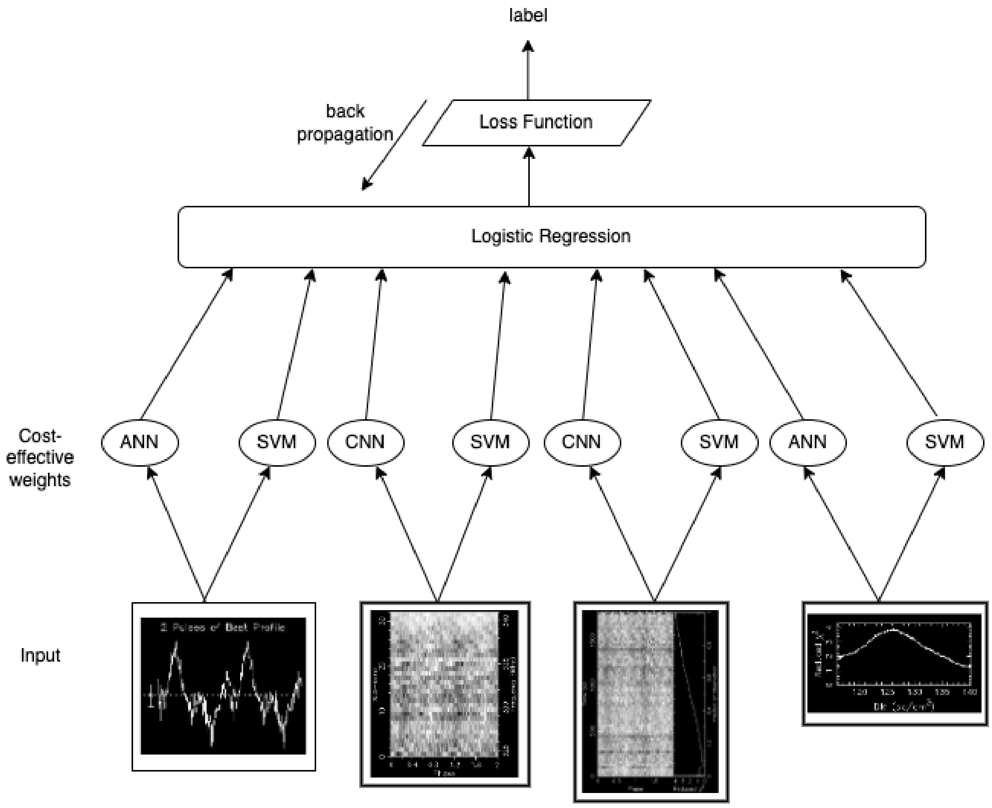

Figure 2: An illustration of model's basis

3 Data and Methodology

3.1 Data Collection and Preprocessing

The data utilized in this study was collected from the Giant Metrewave Radio Telescope (GMRT). To prepare the data for analysis, we employed the Sigrpoc-3.7 tool (Mott (2003)), which converted the raw files into filterbank format. Subsequently, we harnessed the capabilities of PRESTO, a software package designed for pulsar signal analysis. By applying standard search methods (Ransom (2001)) , we generated a set of 18,352 candidates. This entire data processing pipeline was custom-made to meet the specific requirements of our analysis method.

3.2 Data Labeling

The next crucial step involved manually reviewing each candidate and assigning labels based on specific criteria. These criteria allowed us to differentiate between pulsar candidates and non-pulsar candidates, a meticulous process that is central to the accuracy of our study.

3.3 Dataset Generation for Deep Learning

With labeled data in hand, we created a dataset following the ImageNet format. This format is tailored to seamlessly integrate with our machine learning model, ensuring that it can directly ingest and process the data.

3.4 Conclusion

In conclusion, the study of pulsars holds great promise for advancing our understanding of the universe, serving as precision cosmic clocks, tools for interstellar exploration, and key instruments for space navigation. To fully harness this potential, astronomers are adopting a multi-modal approach that integrates deep learning and traditional machine learning techniques, combining CNNs and logistic regression to improve pulsar candidate detection. The carefully curated dataset sourced from the GMRT, formatted in the ImageNet style, provides a robust foundation for this research. This endeavor aims to unlock the full potential of pulsar candidate identification, offering enhanced accuracy and efficiency. The fusion of advanced deep learning with traditional machine learning creates a powerful tool for groundbreaking discoveries in the study of pulsars and the broader universe, positioning this research at the forefront of astronomical advancements with the promise of refining our understanding of pulsars and catalyzing new revelations in astrophysics.

4 Future Directions

In summary, our ongoing research is a crucial step toward improving pulsar candidate identification using deep learning. Future directions include developing advanced model variants, optimizing model performance, addressing class imbalance, enhancing model interpretability, handling noisy data, exploring data augmentation, and validating our models with new data to advance astrophysical research.

References

Mott, A. J., F. P. C. (2003). Pulsar searches in globular clusters at 327 mhz. *American Astronomical Society Meeting 203*.

Ransom, S. M. (2001). *New Search Techniques for Binary Pulsars*. Ph.d. thesis, Harvard University. Publication Number: AAT3028434, ISBN: 9780493408415.

Zhu, W., Berndsen, A., Madsen, E., Tan, M., Stairs, I., Brazier, A., Lazarus, P., Lynch, R., Scholz, P., Stovall, K., Ransom, S., Banaszak, S., Biwer, C., Cohen, S., Dartez, L., Flanigan, J., Lunsford, G., Martinez, J., Mata, A., and Verkataraman, A. (2014). Searching for pulsars using image pattern recognition. *The Astrophysical Journal*, 781(2):117.