Deep Learning to Reconstruct Gas Skymaps for Dark Matter Detection

Alexander Shmakov, Christopher M. Karwin*, Mohammadamin Tavakoli*, Simona Murgia, Pierre Baldi*   ** = Equal Contribution*

Background

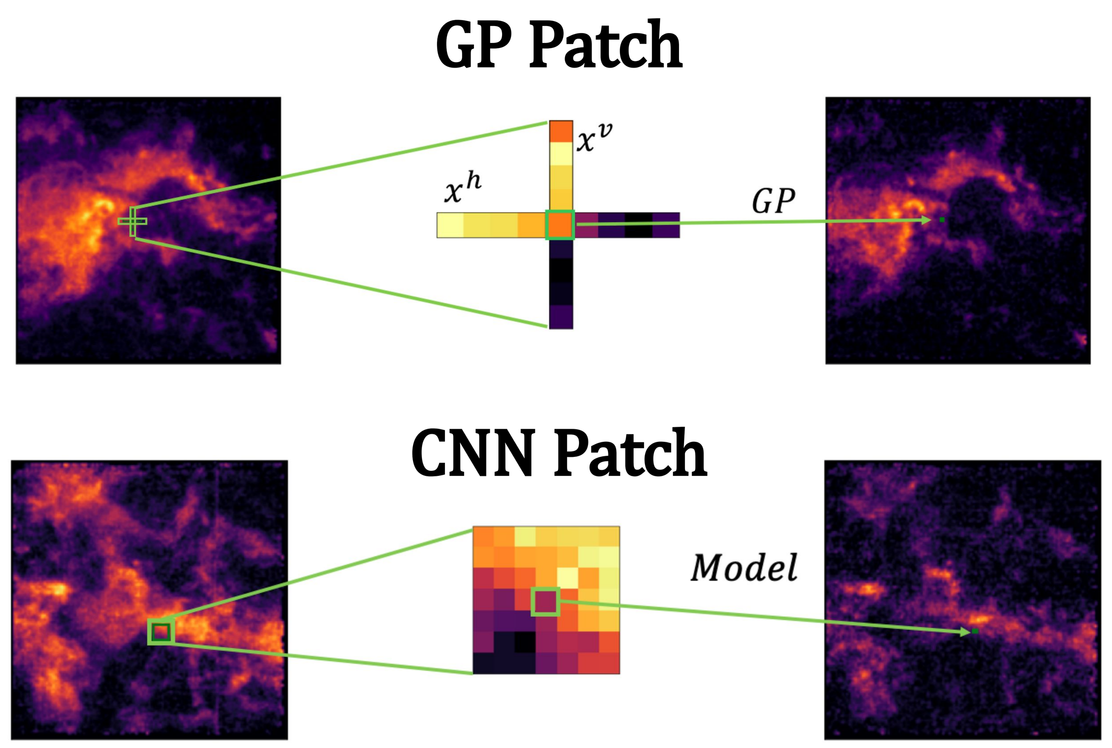
- The Fermi Large Area Telescope (LAT) has **detected excess gamma-rays** towards the Galactic center[1].
- Many possible explanations for this signal, including **annihilating dark matter** in the central Milky Way[1].
- However, we first need to first remove all **explainable sources** of gamma radiation to account for background signal.
- One important indicator for gamma radiation emitted by cosmic rays interacting with dense interstellar gas is one of the isotopologues of carbon monoxide: ^{13}CO.

Overview of Method

- **However**, ^{13}CO is typically difficult to measure directly, but we can easily measure the more common isotopologue: ^{12}CO.
- Maybe we can learn a mapping between these two isotopologues.

$$^{12}\text{CO} \rightarrow {}^{13}\text{CO}$$

- Recently, the Mopra Southern Galactic Plane CO Survey[2] has provided a dataset with both isotopologues in the same region.
- MOPRA provides 50, 1° x 1° x 17 bins, image cubes of the galactic center. **Not enough for a traditional image-to-image modeling!**
- Since concentration is a local process, We **split these images into patches** and train on each of them individually!
- We train a **Matern-kernel Gaussian process** and two variations of a **convolutional neural network** in order to learn this mapping.



References

1. Dan Hooper and Lisa Goodenough. *Dark Matter Annihilation in The Galactic Center As Seen by the Fermi Gamma Ray Space Telescope.* Phys. Lett. B, 697:412–428, 2011.
2. Michael G Burton, Catherine Braiding, Christian Glueck, Paul Goldsmith, Jarryd Hawkes, David J Hollenbach, Craig Kulesa, Christopher L Martin, Jorge L Pineda, Gavin Rowell, et al. *The Mopra Southern Galactic Plane CO Survey.* Publications of the Astronomical Society of Australia, 30, 2013.

Gaussian Process

- Matern-kernel Gaussian process with a whitening kernel to account for the noise in the data.
- Train row-wise and column-wise gaussian process independently at each velocity bin and combine into a single prediction.

$$o_{[i,j]} = \frac{y^h_{[i,j]} + y^v_{[i,j]}}{2}$$

CNN

- We include a stack of convolution layers at the front that evaluate the entire patch. Each layer in the stack is only responsible for a single velocity bin.
- Afterwards, we add hidden feed-forward layers that operate on the resulting latent vector, sharing weights between velocity bins.
- A final fully-connected layer produces the scalar ^{13}CO estimate.

CNN Loss

- MOPRA contains many regions of space with **sparse concentrations** of ^{13}CO and also several bright **hot-spots** with incredibly high concentration.
- We need a loss which **works across many orders of magnitude**.
- We compare two loss functions: **Weighted MAE** & **Poisson NLL**

$$\gamma_t = \left(\sigma - \frac{\sigma}{1 + \exp\left(-10t + 5\right)}\right)\left(\text{Rescale}(T_i) + 1\right)$$

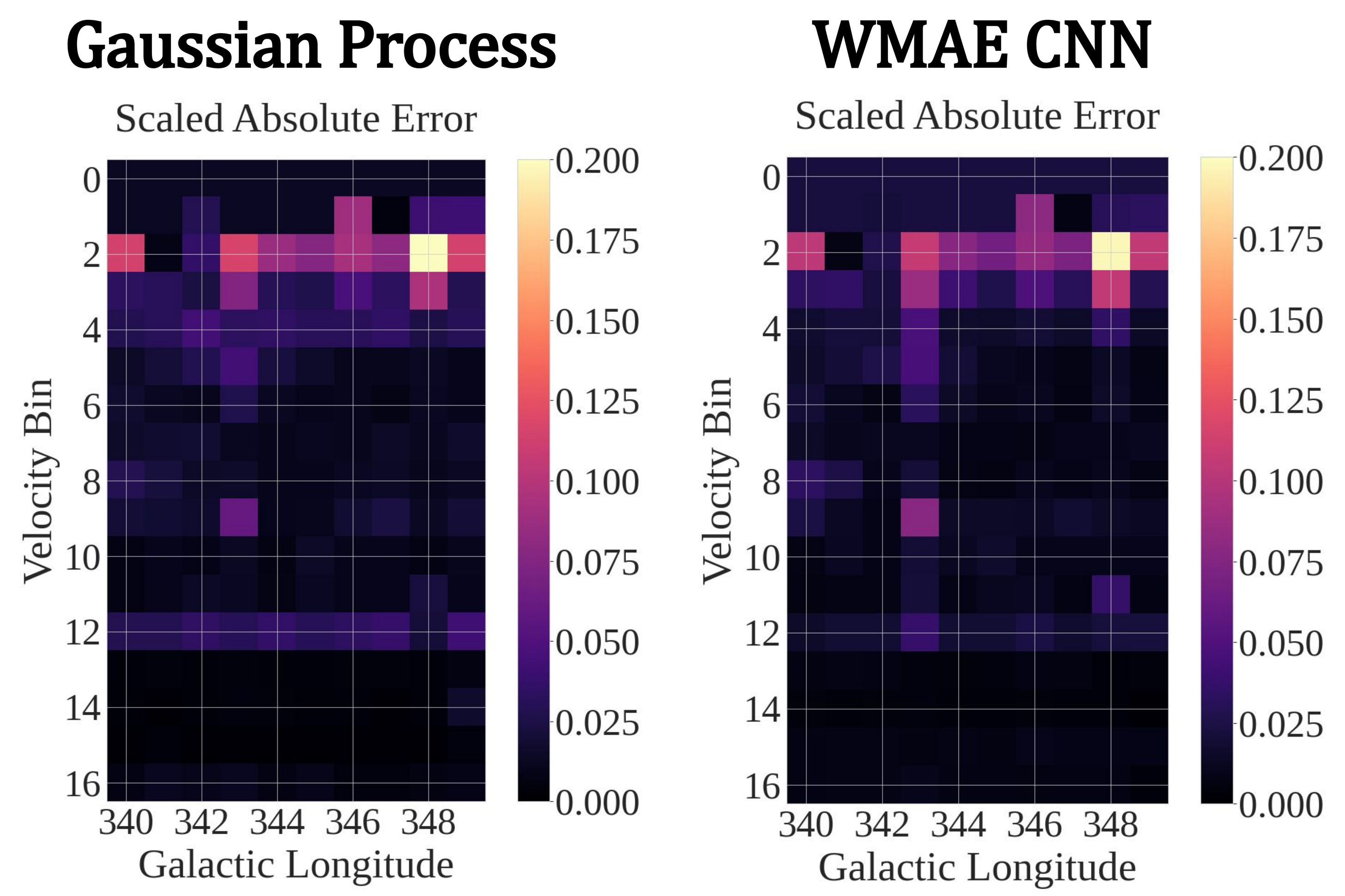
$$\mathcal{L}_{\text{WMAE}} = \frac{1}{N}\sum_{i=1}^{N} \gamma_t \left|\left(f(I_i) - T_i\right|\right.$$

$$\mathcal{L}_{\text{Poisson}} = \frac{1}{N}\sum_{i=1}^{N} f(I_i) - T_i \log f(I_i)$$

Results

Cannot use percent-error directly due to data variability. **However**, **bright regions are the most important**. We evaluate how well we model ^{13}CO *w.r.t* these bright regions: **Scaled Absolute Error**

$$\frac{1}{N}\sum_{n=1}^{N} \frac{|f(I_n) - T_n|}{\max_n |T_n|}$$



Locality

Was our assumption that we can train on individual patches a valid one? We compare model performance at different patch sizes.