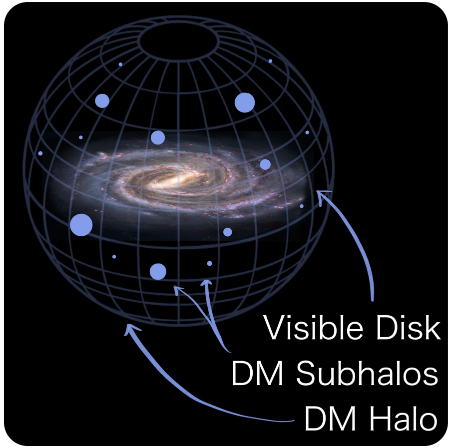
Sep 11th 2023

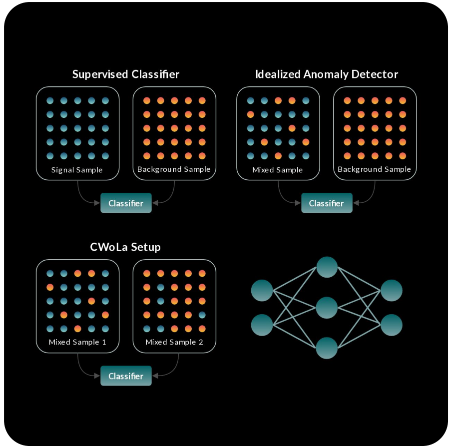
Searches for Exotic Objects among Fermi-LAT γ-Ray Sources with (Weakly) Supervised Machine Learning

*Michael Krämer[1], Silvia Manconi[2], **Kathrin Nippel**[1]*

[1] TTK RWTH Aachen, [2]LAPTh, CNRS

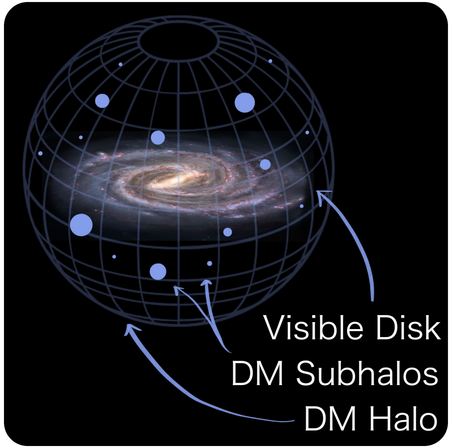
Overview

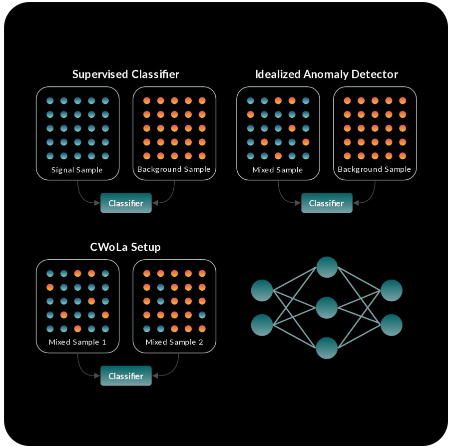

Visible Disk
DM Subhalos
DM Halo


Supervised Classifier
Signal Sample
Background Sample
Classifier
Idealized Anomaly Detector
Mixed Sample
Background Sample
Classifier
CWoLa Setup
Mixed Sample 1
Mixed Sample 2
Classifier

Overview

Search for signals of new physics

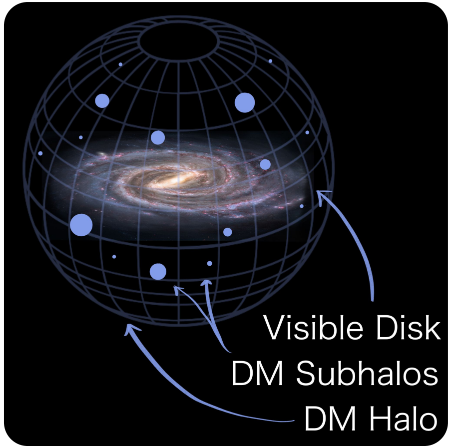


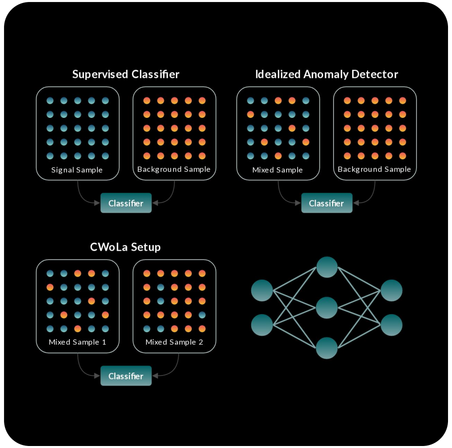


- Indirect dark matter search
- Potential γ-ray signal from WIMP annihilation
- DM subhalos are interesting candidate

Overview

Find exotic signals among Fermi-LAT observation

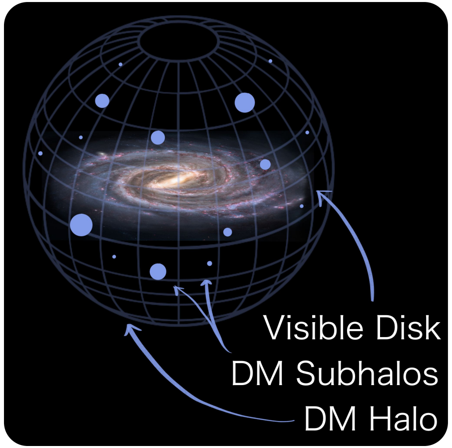




- Indirect dark matter search
- Potential γ-ray signal from WIMP annihilation
- DM subhalos are interesting candidate

- Use abundant γ-ray point source catalog (4FGL)
- Compare known astrophysical objects to unclassified objects

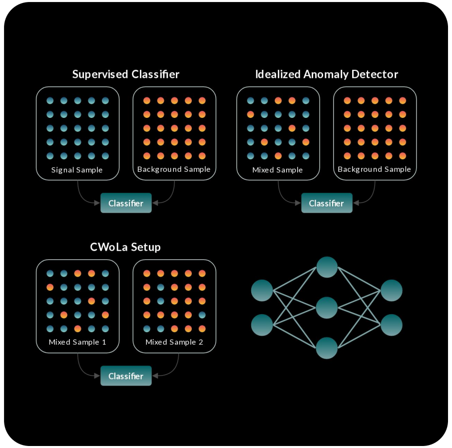
Overview



- Indirect dark matter search
- Potential γ-ray signal from WIMP annihilation
- DM subhalos are interesting candidate

- Use abundant γ-ray point source catalog (4FGL)
- Compare known astrophysical objects to unclassified objects

Machine Learning Approach



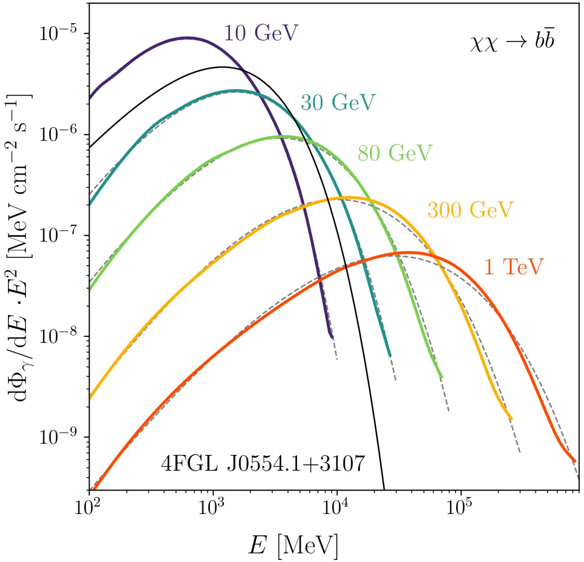
- Supervised: Model expected signal, take uncertainties into account
- Weakly supervised: Model agnostic approach

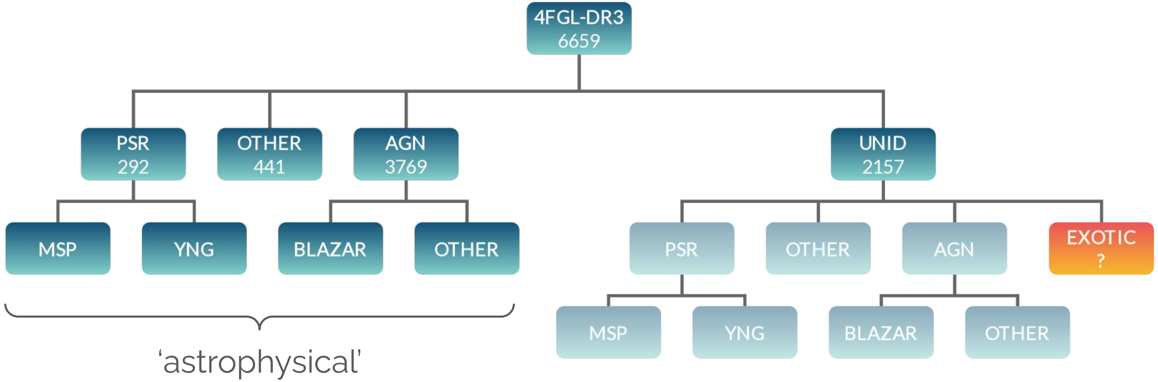
Context

- Dark matter annihilation can lead to a photon flux, detectable in the γ-ray band

$$\phi_\gamma = \frac{\partial \Phi_\gamma}{\partial E}(E, \Delta\Omega) = \frac{\langle \sigma v \rangle}{8\pi m_{\mathrm{DM}}^2}\, \mathcal{J}(\Delta\Omega)\, \frac{\partial N_\gamma^i}{\partial E}(E)$$

- Signal from objects like dark matter subhalos could already have been detected with Fermi-LAT

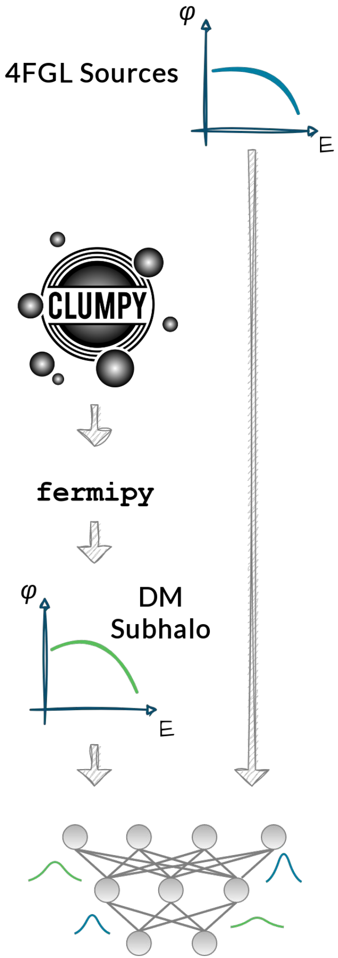




Simulation of Dark Matter Annihilation Spectra

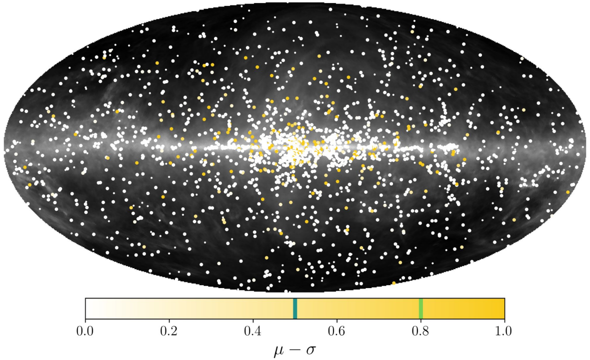
First strategy:
Compare observed, unlabelled objects with modeled subhalo signals to constrain annihilation.

1. Create realistic set of subhalo simulations
2. Assess detectability
   ➜ strongly depends on specific dark matter model
3. Look for subhalo-like spectra among unclassified sources with Bayesian machine learning
   ➜ Deep neural network with probability distributions to describe trainable parameters
   ➜ Output follows probability distribution
   ➜ Result: Prediction μ and uncertainty σ

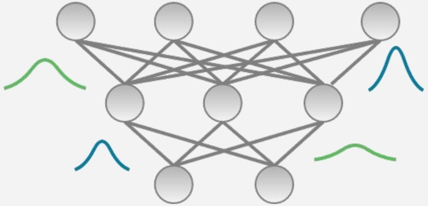


Based on arxiv:2304.00032

Supervised Classification with Bayesian NNs



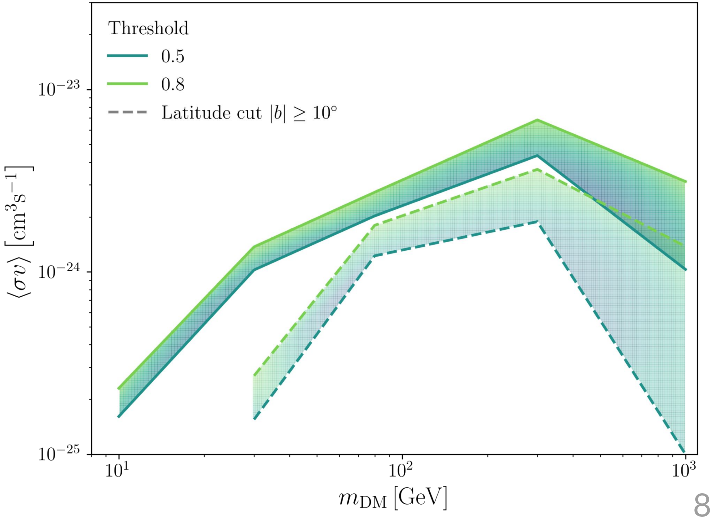
- BNN trained on astrophysical vs subhalo classification task can give predictions on UNID sources
- Get prediction mean and uncertainty on each object. Depending on threshold a source can be deemed *candidate*



- Compare number of sources *not* classified as astrophysical origin with expected detectability of subhalo-like objects:

  ➜ Place conservative limits on annihilation cross-section

- Mass dependence due to detectability and distinguiability



Classification Without Labels (CWoLa 🐨)

➜ Weakly supervised machine learning method

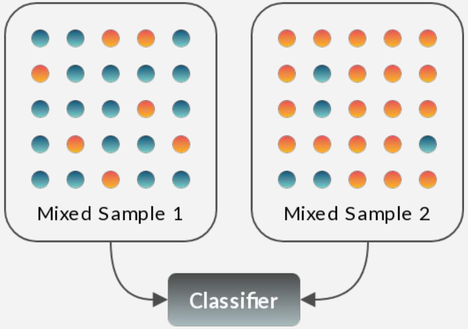
Argument:

- An optimal classifier trained to distinguish the mixed samples is also optimal to distinguish signal from background objects
    - Optimal classifier is given by the likelihood ratio, this relates to:

$$L_{M_1/M_2} = \frac{p_{M_1}}{p_{M_2}} = \frac{f_1 p_S + (1-f_1) p_B}{f_2 p_S + (1-f_2) p_B} = \frac{f_1 L_{S/B} + (1-f_1)}{f_2 L_{S/B} + (1-f_2)}$$

    - Training a classifier to maximize L_{M_1/M_2} yields the optimal classifier also to discriminate signal and background if $f_1 > f_2$.

CWoLa Setup



Following ArXiv:1708.02949

Classification Without Labels (CWoLa)

➜ Weakly supervised machine learning method

Argument:

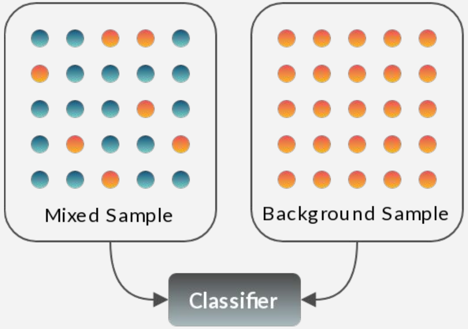
- An optimal classifier trained to distinguish the mixed samples is also optimal to distinguish signal from background objects
  - Optimal classifier is given by the likelihood ratio, this relates to:

$$L_{M_1/M_2} = \frac{p_{M_1}}{p_{M_2}} = \frac{f_1 p_S + (1-f_1) p_B}{f_2 p_S + (1-f_2) p_B} = \frac{f_1 L_{S/B} + (1-f_1)}{f_2 L_{S/B} + (1-f_2)}$$

  - Training a classifier to maximize L_{M_1/M_2} yields the optimal classifier also to discriminate signal and background if $f_1 > f_2$.

- In our approach: signal ≙ exotic source, background ≙ astrophysical source *Idealized setup* due to pure background sample
  - Setup test approach where *f* can be controlled and where we can compare to supervised classification

Idealized Anomaly Detector

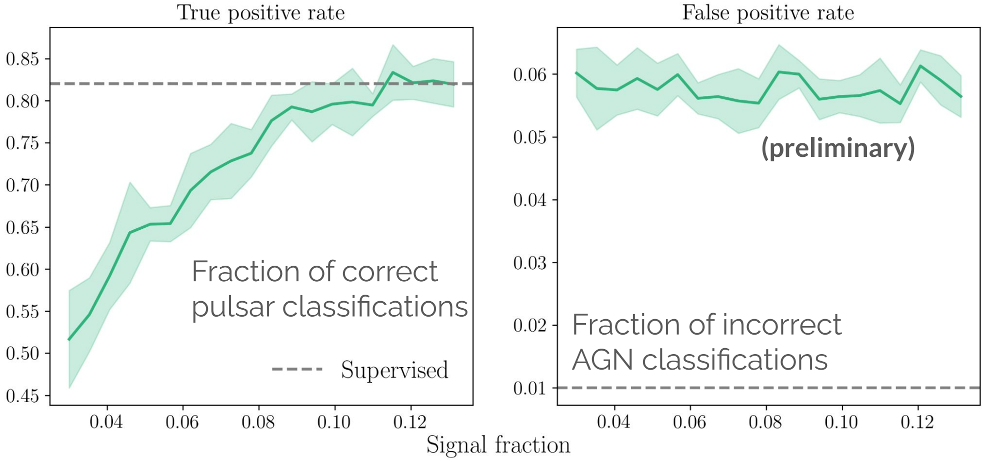


Following ArXiv:1708.02949

Classification Without Labels (CWoLa)

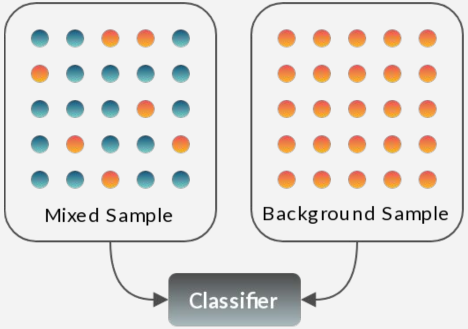
Test case:

Signal $\leftrightarrow$ Pulsar spectra,
Background $\leftrightarrow$ AGN spectra (+ data augmentation)



Fraction of pulsars in mixed sample

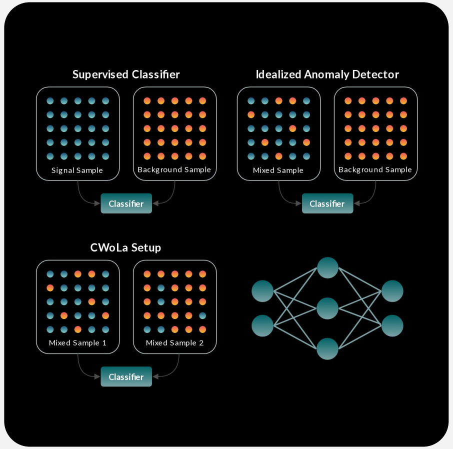
Idealized Anomaly Detector



Result:

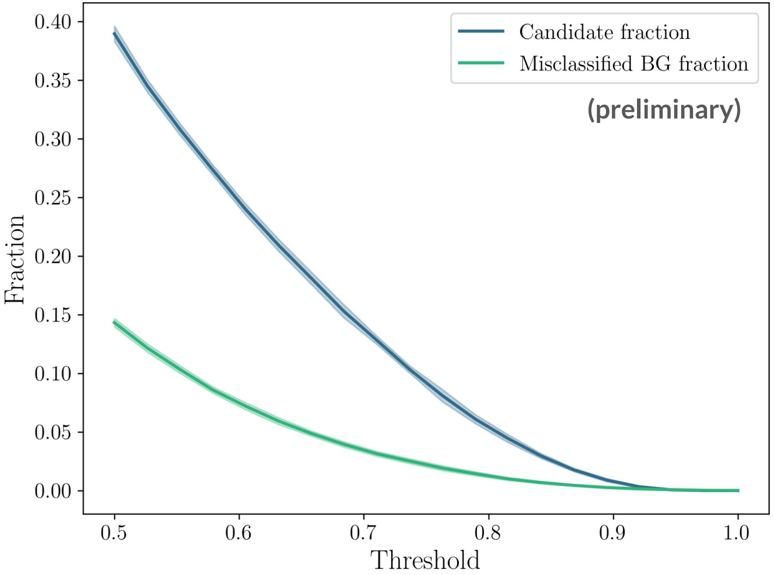
Method yields promising results, but is overall limited by small sample size

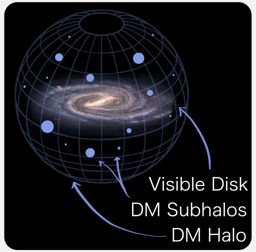
Results for UNID classification

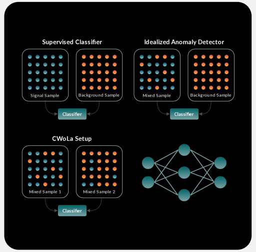


- Supervised classification can help define interesting candidates among UNID samples with corresponding prediction uncertainty

- Weakly supervised classification applied to UNID vs astrophysical classification task gives weaker hints for exotic objects, but is model independent

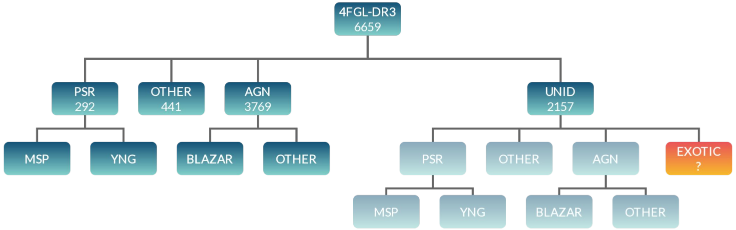




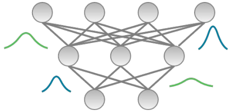


Conclusion and Outlook

Two machine learning approaches to learn about the abundance of exotic objects in Fermi-LAT data:



1. Supervised: Model dark matter subhalo spectra from WIMP annihilation and subhalo population model, robust classification with Bayesian machine learning



2. Weakly supervised: Model independent approach based on *CWoLa*, intriguing first results that we aim to strengthen with more testing and more advanced methods