Deep Unsupervised Domain Adaptation for the Cherenkov Telescope Array

Michaël Dell'aiera*

*Under the supervision of Thomas Vuillaume** and Alexandre Benoit***.*

European Escape program (824064)*
Laboratoire d'Annecy de Physique des Particules (LAPP)**
Laboratoire d'Info, Systèmes, Traitement de l'Info et de la Connaissance (LISTIC)***

September 27, 2022

Presentation Outline

Gamma-ray astronomy

Study of the **high-energy gamma sources** in the Universe.

Inverse problem resolution : Given the **telescope observations**, how to recover the **Energy**, **Direction** and **Type** of the incoming particle ?

- Single telescope analysis (LST-1, CTA project)
- Particle classification : Gamma or Proton (maybe Electron ?)

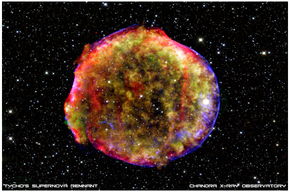


Figure 1: Supernova. Source [1].

Figure 2: Black hole. Source [1].

Figure 3: Dark matter. Source [1].

Principle of detection

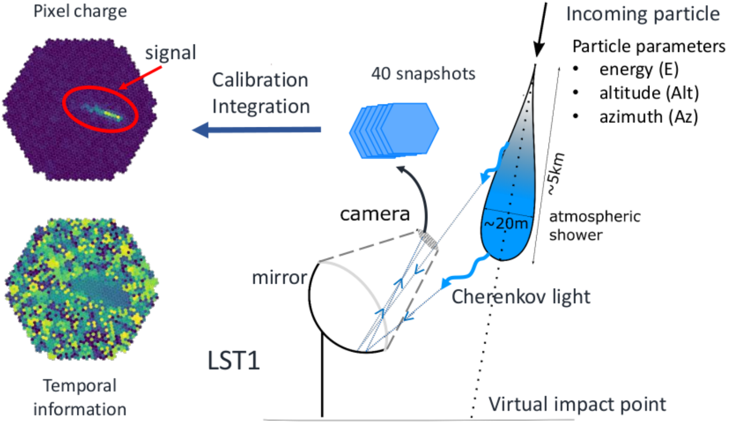


Figure 4: Principle of gamma particle detection. Source [8].

Workflow

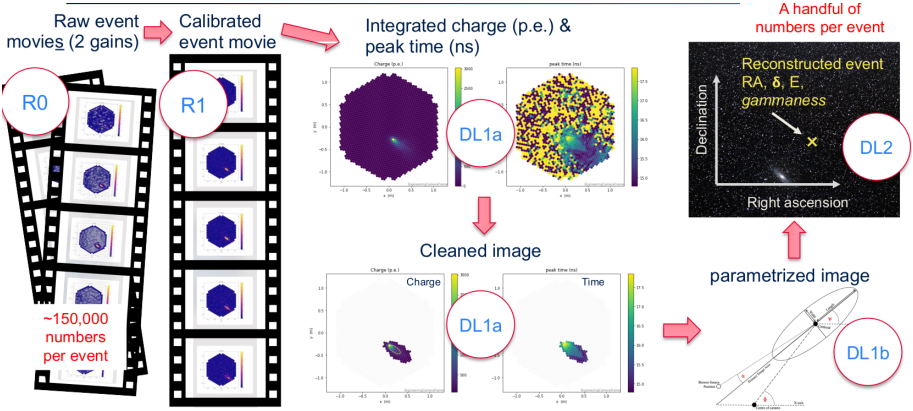


Figure 5: Particle detection workflow (Hillas). Source : LST Analysis School.

Workflow

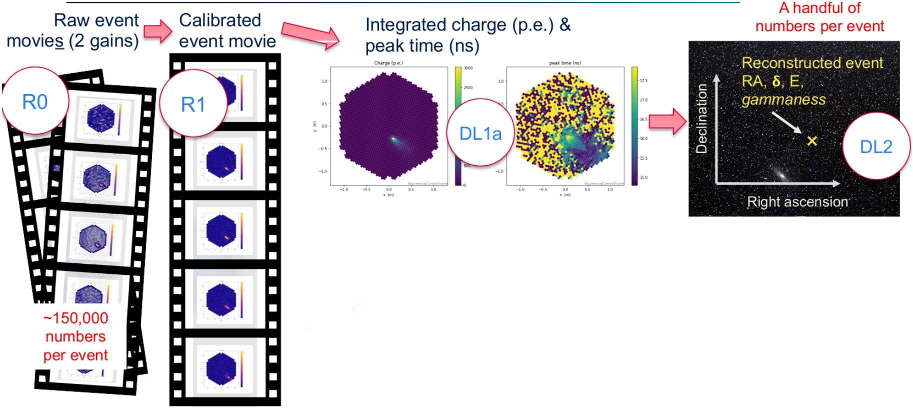


Figure 6: Particle detection workflow (γ-PhysNet).

γ-PhysNet

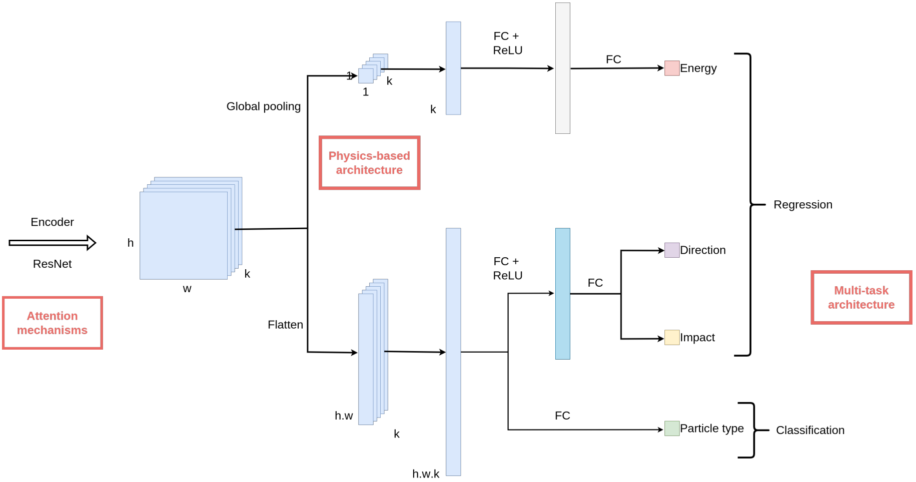


Figure 7: γ-PhysNet architecture. Source [5].

Results on MC simulations

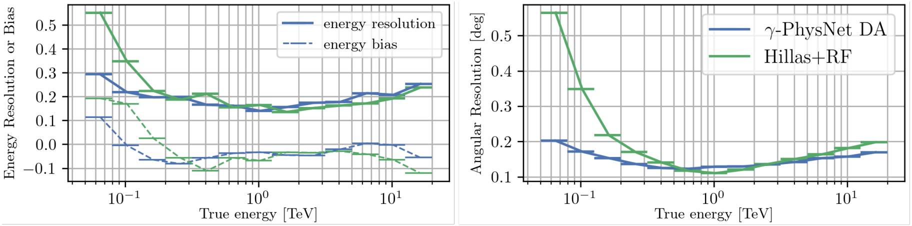


Figure 8: Comparison of different IRFs between γ-PhysNet and Hillas+RF on simulated data. Left : Energy resolution as a function of the true energy. Right : Angular resolution as a function of the true energy. In both cases, **lower is better**. Source [8].

Results on MC simulations

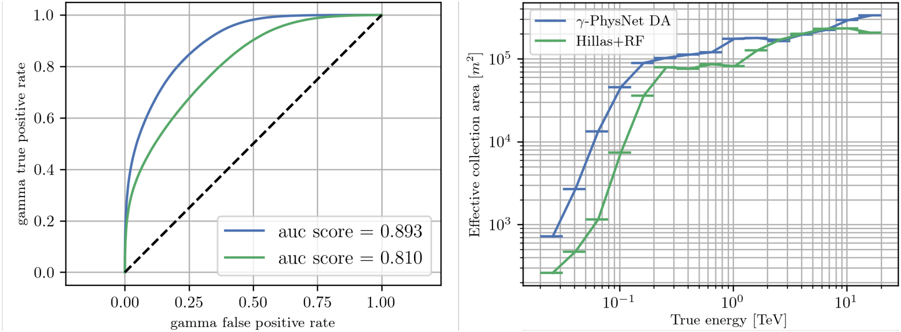


Figure 9: Comparison of different IRFs between γ-PhysNet and Hillas+RF on simulated data. Left : ROC curves of the particle classification. Right : Effective collection area as a function of the true energy. In both cases, **higher is better**. Source [8].

Matching real data

- Unobtainable faultless labelled real data

  → Training fully relying on simulations

- Simulations are *close approximations* of the reality

  → Add Poisson noise ($P(\lambda)$) to match the NSB distributions

Training	**Inference**
MC+$P(\lambda)$	Real observations

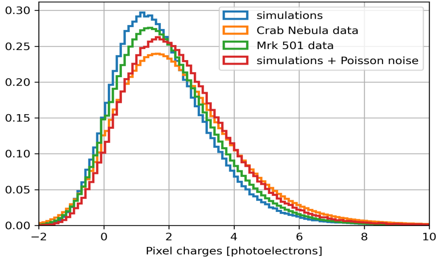


Figure 10: Night Sky Background (NSB) charge distributions for simulated and real data. Source [8].

- $NSB_{MC} \neq NSB_{real}$
- $NSB_{MC} + P(\lambda) \sim NSB_{real}$.

Results on real data

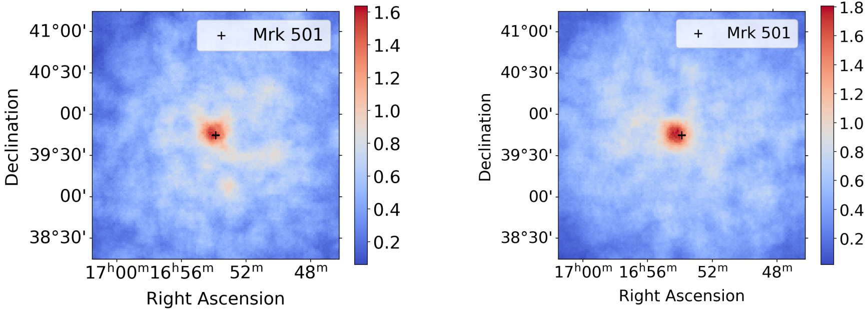


Figure 11: Markarian 501 detection. Left : Hillas + RF. Right : γ-PhysNet. γ-PhysNet detects more events but the area of detection is not centred on the real position. Source [8].

Domain adaptation

Set of algorithms and techniques which aims to reduce domain discrepancies.

- Generalisation of the model ($\neq$ run-wise adaptation)
- Great variety of state-of-the-art methods
- Deep unsupervised domain adaptation
- Promote a domain-invariant representation through domain confusion

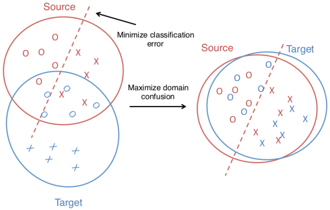


Figure 12: Domain confusion.

Figures of merit

What are the tools at our disposal ?

Confusion matrix

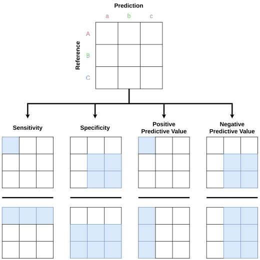


Figure 13

t-SNE

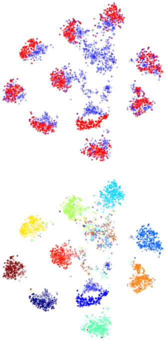


Figure 14

Grad-CAM

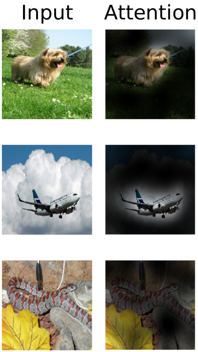


Figure 15

t-SNE

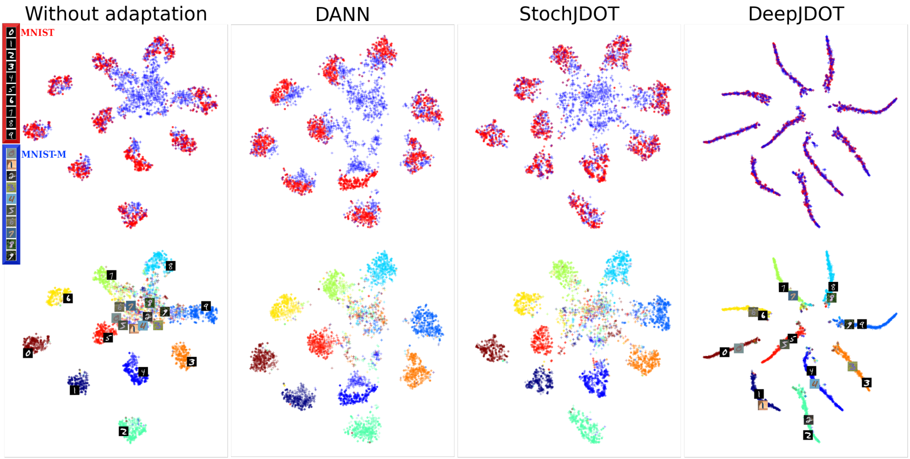


Figure 16: Feature adaptation using 4 distinct domain adaptation methods. Source [2].

γ-PhysNet extended with DANN [4]

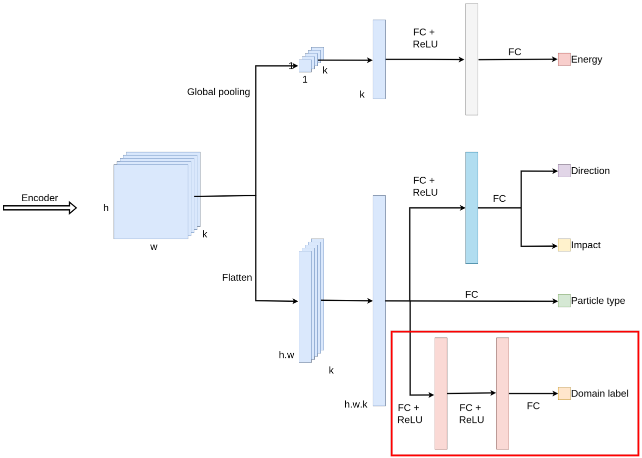


Figure 17: γ-PhysNet coupled with a domain classifier. Inspired from [4].

γ-PhysNet extended with DeepJDOT [2]

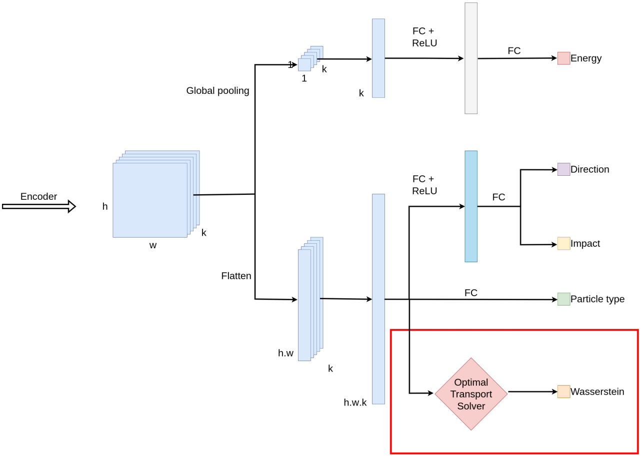
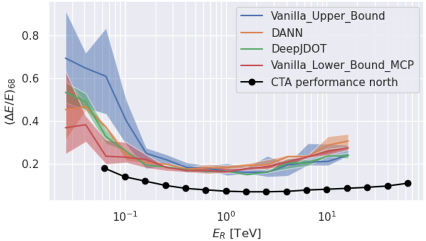


Figure 18: γ-PhysNet coupled with optimal transport. Inspired from [2]

Results on MC simulations



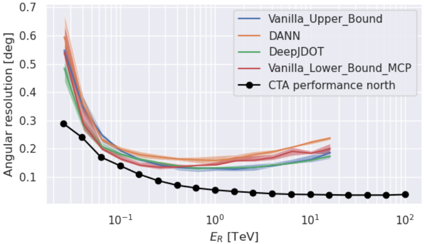


Figure 19: IRFs from DANN and DeepJDOT when trained on MC (Source) and MC+Poisson($\lambda = 0.4$) (Target) then tested on MC+Poisson($\lambda = 0.4$). Left : Energy resolution as a function of the energy. Right : Angular resolution as a function of the energy. In both cases, **lower is better**.

Results on MC simulations

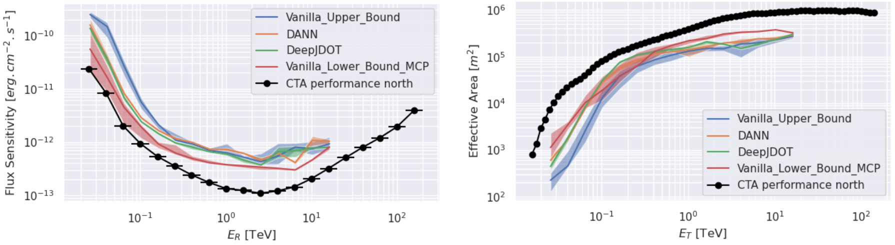


Figure 20: IRFs from DANN and DeepJDOT when trained on MC (Source) and MC+Poisson($\lambda = 0.4$) (Target) then tested on MC+Poisson($\lambda = 0.4$). Left : Flux sensitivity as a function of the energy. **Lower is better**. Right : Effective area as a function of the energy. **Higher is better**.

Conclusion & Perspectives

- Ultimate goal: Create an event map to localise gamma-ray sources
- Currently trying to reduce the discrepancy between MC and real data

- Working in parallel on:
  - Replace MC+P(λ) with real data (Markarian 501, Crab, ...)
  - The pointing direction between training data and observations
  - Introduction of physics knowledge (Physics-based deep learning)
  - Upgrading DANN [4] $\rightarrow$ WDGRL [6]

- Future challenges:
  - Stereoscopy (Federated learning [9], ...)
  - Neural network explainability (Information bottleneck [7], ..)
  - Continual learning [3]

γ − learn

GammaLearn

- All info and contact: https://gammalearn.pages.in2p3.fr/pages/

- Acknowledgments

- This project is supported by the facilities offered by the **Univ. Savoie Mont Blanc - CNRS/IN2P3 MUST computing center**
- This project was granted access to the HPC resources of **IDRIS under the allocation 2020-AD011011577 made by GENCI**
- This project is supported by the computing and data processing ressources from the **CNRS/IN2P3 Computing Center (Lyon - France)**
- We gratefully acknowledge the support of the **NVIDIA Corporation** with the donation of one NVIDIA P6000 GPU for this research.
- We gratefully acknowledge financial support from the agencies and organizations listed here: **https://www.cta-observatory.org/consortium_acknowledgment**.
- This project has received funding from the ***European Union's Horizon 2020 research and innovation programme*** under grant agreement **No 653477**
- This project has received funding from the ***European Union's Horizon 2020 research and innovation programme*** grant agreement **No 824064**

References

[1] *cta-observatory*. https://www.cta-observatory.org/.

[2] Bharath Bhushan Damodaran et al. "DeepJDOT: Deep Joint Distribution Optimal Transport for Unsupervised Domain Adaptation". In: (2018). DOI: 10.48550/ARXIV.1803.10081. URL: https://arxiv.org/abs/1803.10081.

[3] Matthias Delange et al. "A continual learning survey: Defying forgetting in classification tasks". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021), pp. 1–1. DOI: 10.1109/TPAMI.2021.3057446.

[4] Yaroslav Ganin et al. *Domain-Adversarial Training of Neural Networks*. 2016. arXiv: 1505.07818 [stat.ML].

List of References

[5] Mikaël Jacquemont. "Cherenkov Image Analysis with Deep Multi-Task Learning from Single-Telescope Data". Theses. Université Savoie Mont Blanc, Nov. 2020. URL: https://hal.archives-ouvertes.fr/tel-03590369.

[6] Jian Shen et al. *Wasserstein Distance Guided Representation Learning for Domain Adaptation*. 2018. arXiv: 1707.01217 [stat.ML].

[7] Naftali Tishby and Noga Zaslavsky. *Deep Learning and the Information Bottleneck Principle*. 2015. arXiv: 1503.02406 [cs.LG].

[8] Thomas Vuillaume et al. "Analysis of the Cherenkov Telescope Array first Large-Sized Telescope real data using convolutional neural networks". In: *arXiv preprint arXiv:2108.04130* (2021).

[9] Qiang Yang et al. 2019.

Thank you for your attention!