

Application of Machine Learning in Astroparticle Physics

ICHEPAP 2023

Saha Institute for Nuclear Physics

Daniel Nieto (d.nieto@ucm.es) Institute for Particle and Cosmos Physics IPARCOS-UCM





- Machine learning in a nutshell
- Enhancing IACTs with machine learning
- Applications to other astroparticle physics experiments





Artificial intelligence?

Problems in AI:

Study of *intelligent agents*: systems that perceive their environment and take actions that maximize their chance to successfully achieving their goals (e.g. solving a problem).

Machine Learning

- System acquires its own knowledge learning from examples
- System achieves its goal without utilizing explicit rules
- Abstract and formal:
 computers, Solution humans
 Knowledge-based approach



Intuitive, hard to formalize:
 humans, so computers
 Concept capture and generalization



Machine Learning in a nutshell







D. Nieto



Machine Learning strategies





towardsdatascience.com













https://xkcd.com/1838/







cs231n.github.io







- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ -ray collection area (~10⁵ m²)
- Large background from charged CR
 - Partly irreducible (e⁻/e⁺, single-EM, with current methods)
- \circ Energy window: tens GeV tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction







- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ -ray collection area (~10⁵ m²)
- Large background from charged CR
 - Partly irreducible (e⁻/e⁺, single-EM, with current methods)
- Energy window: tens GeV tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction























Output: event type, energy, arrival direction



Input: observed events



ML method: o Random Forest (RF)

• Applied to:

Current generation of IACTs: classic ML

.

- o Background rejection
- o Arrival direction



Aleksic et al., A&A 524 A77 (2010)











- ML method:
 - Boosted Decision Trees (BDT)
- Applied to:

Background rejection



Current generation of IACTs: classic ML





Output: event type, energy, arrival direction



Input: observed events

Current generation of IACTs: classic ML



• ML method:

Boosted Decision Trees (BDT)

• Applied to:

Background rejection





(Results for H.E.S.S. I only)





Convolution

Kernel

Guo et al.

Outputs

Output: event type, energy, incoming direction Convolutional Neural Network (CNN) INPUT feature maps feature maps feature maps feature maps OUTPUT Output e 28x28 4@24x24 4@12x12 12@8x8 12@4x4 10@1x1 v e n Mapping from Output Output t features Additional e Mapping from Mapping from layers of more LeCunn et al. Output features features abstract features DL capable of *extracting* and mapping image features automatically with 0 0 unprecedented classification accuracy. Hyper-active CS research field constantly n Hand-Hand-Simple S designed designed Features improving features program features Many HEP/Astro experiments already exploring/utilizing the technique (LIGO, 0 LHC, MicroBooNe, NOVa, etc...) u C Method: Input Input Use deep learning to reconstruct IACT events from non-parameterized images 0 Deep Performance enhancement -> better sensitivity Classic learning Rule-based machine 0 systems Representation learning learning n E.g.: RF & BDT But there are risk...

MC reliability (e.g. network selecting some features from your MC not present in 0 real data)

Input: observed events







www.cta-observatory.org

Science with CTA, arXiv:1709.07997



Challenges for machine learning from IACT data



Stereoscopy:

Stereoscopic view of the extended air showers
Compact "videos" rather than single snapshots
Events effectively recorded in 4D!

CREDIT: DESY/Milde Science Communication





• Heterogeneity of instruments:

Camera images courtesy of T. Vuillaume







• Heterogeneity of instruments:



D. Nieto





• Final metrics are far from trivial and entangled

Flux sensitivity

Energy resolution





Credit: <u>www.cta-observatory.org</u>











- Single telescope
- Square pixels
- Only signal charge (no timing)
- Single task: classification



Medium energies (0.3 TeV < E < 1 TeV)



| AUC |
|-----|
|-----|

| Model/Energy | Low E. | Med. E. | High E. |
|--------------|--------|---------|---------|
| InceptionV3 | 84.7% | 91.1% | 92.0% |
| ResNet50 | 84.8% | 91.4% | 90.2% |





- High-level Python package for using deep learning for IACT event reconstruction
- Configuration-file-based workflow and installation with conda drive reproducible training and prediction
- Supports any TensorFlow model that obeys a generic signature
- Open source on GitHub:

https://github.com/ctlearn-project/ctlearn https://pos.sissa.it/358/752 DOI 10.5281/zenodo.3345947 (Latest release: CTLearn v0.6.2)



<u>Core developers</u> Tjark Miener, DN (I**PARCOS-UCM**) Ari Brill, Qi Feng (Columbia) Bryan Kim (UCLA, now at Stanford) (See contributors <u>here</u>)





Tackling the hexagonal-pixel challenge



earn

Image mapping (preprocessing)





https://github.com/ctlearn-project/



 \checkmark Angles and distances preserved





Hexagonal convolution CRAPP T. Vuillaume, O M. Jaquemont, et al. learn Convolution https://github.com/IndexedConv Index matrix Axial addressing system Wх Convolution kernel 17 18 Image stored as a vector Pooling Index matrix 6 7 2 3 Rebuild index matrix (M. Jacquemont et al. 2019)





• Comparison of methods for classification task



D. Nieto et al. PoS(ICRC2019)753



CTLearn: single-telescope full-event reconstruction











٠

٠

٠

•

Vlearn

Different models

CTLearn: crosschecking results



















T. Miener et al., PoS(ICRC2021) 730



CTLearn: multiple-telescope full-event reconstruction













T. Miener et al. 2021 (ADASS XXXI)







T. Miener et al. 2021 (ADASS XXXI)



Dreaming of IACTs







Dreaming of IACTs








• Auxiliary conditional generative adversarial networks (AC-GANs)







IceCube





- Graph Neural Networks for low-energy events
- Classification: signal/bkg, neutrino types
- Signal eff. +18%, bkg suppression /8
- Energy, arrival direction ~20% improv
- Processing rate 2x IceCube trigger rate (GPU)

JINST 17 P11003 (2022)







AUGER OBSERVATORY

- X_{max} from Fluorescence Det. (~15% duty cycle)→ X_{max} from Surface Det. (~99% duty cycle) (Xmax: air shower maximum, nuclear mass from UHCR)
- CNN with hexagonal convolution kernels



JINST 16 P07019 (2021)







- Attention-based CNN architecture
- Determination of arrival direction
- Best performing than standard analysis on simulation, but not on real observations yet







- o Current-generation IACTs have enhanced their performances through ML
- Next-gen (even current-gen!) IACT may profit from latest developments in ML
- o Ongoing efforts to exploit deep learning as an event reconstruction method for IACTs
 - Full-event reconstruction over simulated IACT events demonstrated
 - Application to real observations works!
 - Working on optimizing architectures & multi-task learning
 - Using AC-GANs as pseudosimulators
 - Tackling the real-data problem







The research here presented has been partially supported by the former Spanish Ministry of Economy, Industry, and Competitiveness / ERDF grants FPA2015-73913-JIN and FPA2017-82729-C6-3-R, the Spanish Ministry of Science and Innovation grant PID2019-104114RB-C32, NSF awards PHY-1229205, 1229792, and 1607491, and the European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures funded by the European Union's Horizon 2020 research and innovation program under Grant Agreement no. 824064. The authors acknowledge support from Google LLC through the Google Summer of Code program and NVIDIA Corporation with the donation of a Titan X Pascal GPU used for part of this research.











Deep Convolutional Neural Networks





<u>cs231n.github.io</u>



Gamma-ray detectors













Conventional non-thermal emitters & CR accelerators





Gamma-ray Binaries



Pulsars



Gamma-ray Bursts



Compact-object mergers





Supernova Remnants

Pulsar Wind Nebulae



Starburst Galaxies

Active Galactic Nuclei



Multimessenger astronomy









Dark matter searches





Lorentz invariance





D. Nieto



Particle showers produced in Earth's atmosphere by gamma-ray, proton, and carbon-13

- Initial particle energy: 400 GeV
- Animation time: Shower reaching ground
- Charged particles: Red dots
- Cherenkov light: Blue dots

Visit http://veritas.sao.arizona.edu

©2012 Martin Schroedter VERITAS & Harvard Smithsonian Center for Astrophysics



Imaging atmospheric Cherenkov technique





- Detection of extended air showers using the atmosphere as a calorimeter
- o Huge γ -ray collection area (~10⁵ m²)
- Large background from charged CR
 - Partly irreducible (e⁻/e⁺, single-EM, with current methods)
- o Energy window: tens GeV tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction









Dropout

Dense

Dropout

LSTM

Dropout

DCN





Single-tel model









CNN-RNN model







Gamma/hadron classification

Samples

Samples



D. Nieto et al. PoS(ICRC2019)752





CNN-RNN model









D. Nieto et al. PoS(ICRC2019)752



CTLearn: crosschecking results







- Crosschecking three different implementations
- Same datasets, same cuts
- Different models
- Comparison against standard analysis (RF)





٠

٠

٠

•

Vlearn

Different models

CTLearn: crosschecking results











• Generative adversarial networks (GANs)







• Auxiliary conditional generative adversarial networks (AC-GANs)







• Generation time

GANS

Using a GPU Nvidia 3090:



SIMULATIONS

- ullet ~ 1 s/event
- Each event consists of one image for each detector
- Depends on what is being simulated and the computational capacity.





Next step -> find the best performing model for event reconstruction

The curse of dimensionality haunts us here too!

- Hyperparameter space for deep learning architecture design
 - Number of CNN layers
 - o Kernel size
 - Activation function
 - o Dropout rate
 - Number of FC layers
 - o Batch size
 - o Learning rate
 - o Optimizer
 - 0 ...

- Optimization strategies
 - o Grid searches
 - o Random searches
 - o Bayesian optimization
 - Evolutionary algorithms
 - Reinforcement learning

0 ...





- Framework for hyperparameter optimization of CTLearn models (Although can be adapted to any config-file based DCN framework)
- o Based on Tune: a scalable hyperparameter tuning library
- Supported optimization strategies:
 - Random search
 - Tree Parzen Estimators
 - Gaussian Processes
 Bayesian optimization
 - Genetic Algorithms
 - Parallel optimization (depending on available hardware)

github.com/ctlearn-project/ctlearn_optimizer



ctlearn-optimizer.readthedocs.io













ICHEPAP - SINP - Dec 23











CTLearn Optimizer: some results



mstn

🗕 Ist

- sstc

100









Single_tel & TPE search



Optimized hyperparameters seem to be telescope-type dependent





Single_tel & TPE search: transfer to CNN-RNN

| Hyperparameters | Telescope Type | Validation Accuracy | Validation AUC | Training Time |
|-----------------|----------------|---------------------|----------------|---------------|
| Base | LST | 73.43% | 0.8285 | 0h 41m 22s |
| Optimized | LST | 74.96% | 0.8422 | 0h 46m 53s |
| Base | SSTC | 80.64% | 0.9072 | 1h 51m 5s |
| Optimized | SSTC | 83.49% | 0.9217 | 3h 31m 43s |
| Base | MSTN | 83.10% | 0.9169 | 2h 15m 52s |
| Optimized | MSTN | 84.20% | 0.9313 | 6h 43m 14s |

| Telescope Type | Metric | Improvement |
|----------------|---------------------|-------------|
| LST | Validation Accuracy | 1.53% |
| LST | Validation AUC | 1.37% |
| SSTC | Validation Accuracy | 2.85% |
| SSTC | Validation AUC | 1.45% |
| MSTN | Validation Accuracy | 1.10% |
| MSTN | Validation AUC | 1.44% |





o Multi-task learning



• Tackling the real-data problem

Using GANs to bridge the gap between performances on simulations and observations

o Model optimization

Combine heterogeneous cameras in one model Implement and test deeper models Enable optimization on large GPU clusters

o Invert models to explore pseudo-simulators

o ...





• Event classification task (AUC)





https://arxiv.org/abs/1912.09898





Event classification task (ACC)





https://arxiv.org/abs/1912.09898


The sensitivity of CTA



Sensitivity of CTA: the next-generation γ -ray observatory





CTLearn: crosschecking results







ICHEPAP - SINP - Dec 23