Astro Machine Learning References

July 10, 2018

References

Acciarri et al.: Convolutional neural networks applied to neutrino events in a liquid argon time projection chamber   **1748-0221-12-03-P03011**

R. Acciarri et al. "Convolutional neural networks applied to neutrino events in a liquid argon time projection chamber". In: *Journal of Instrumentation* 12.03 (2017), P03011. URL: http://stacks.iop.org/1748-0221/12/i=03/a=P03011.

Abstract: We present several studies of convolutional neural networks applied to data coming from the MicroBooNE detector, a liquid argon time projection chamber (LArTPC). The algorithms studied include the classification of single particle images, the localization of single particle and neutrino interactions in an image, and the detection of a simulated neutrino event overlaid with cosmic ray backgrounds taken from real detector data. These studies demonstrate the potential of convolutional neural networks for particle identification or event detection on simulated neutrino interactions. We also address technical issues that arise when applying this technique to data from a large LArTPC at or near ground level.

Aurisano et al.: A convolutional neural network neutrino event classifier
1748-0221-11-09-P09001

A. Aurisano et al. "A convolutional neural network neutrino event classifier". In: *Journal of Instrumentation* 11.09 (2016), P09001. URL: http://stacks.iop.org/1748-0221/11/i=09/a=P09001.

Abstract: Convolutional neural networks (CNNs) have been widely applied in the computer vision community to solve complex problems in image recognition and analysis. We describe an application of the CNN technology to the problem of identifying particle interactions in sampling calorimeters used commonly in high energy physics and high energy neutrino physics in particular. Following a discussion of the core concepts of CNNs and recent innovations in CNN architectures related to the field of deep learning, we outline a specific application to the NOvA neutrino detector. This algorithm, CVN (Convolutional Visual Network) identifies neutrino interactions based on their topology without the

need for detailed reconstruction and outperforms algorithms currently in use by the NOvA collaboration.

Baumgartner et al.: Visual Feature Attribution using Wasserstein GANs
<div align="right">DBLP:journals/corr/abs-1711-08998</div>

Christian F. Baumgartner et al. "Visual Feature Attribution using Wasserstein GANs". In: *CoRR* abs/1711.08998 (2017). arXiv: 1711.08998. URL: http://arxiv.org/abs/1711.08998.

Abstract: Attributing the pixels of an input image to a certain category is an important and well-studied problem in computer vision, with applications ranging from weakly supervised localisation to understanding hidden effects in the data. In recent years, approaches based on interpreting a previously trained neural network classifier have become the de facto state-of-the-art and are commonly used on medical as well as natural image datasets. In this paper, we discuss a limitation of these approaches which may lead to only a subset of the category specific features being detected. To address this problem we develop a novel feature attribution technique based on Wasserstein Generative Adversarial Networks (WGAN), which does not suffer from this limitation. We show that our proposed method performs substantially better than the state-of-the-art for visual attribution on a synthetic dataset and on real 3D neuroimaging data from patients with mild cognitive impairment (MCI) and Alzheimer's disease (AD). For AD patients the method produces compellingly realistic disease effect maps which are very close to the observed effects.

Bockermann et al.: Online Analysis of High-Volume Data Streams in Astroparticle Physics
<div align="right">10.1007/978-3-319-23461-8_7</div>

Christian Bockermann et al. "Online Analysis of High-Volume Data Streams in Astroparticle Physics". In: *Machine Learning and Knowledge Discovery in Databases*. Ed. by Albert Bifet et al. Cham: Springer International Publishing, 2015, pp. 100–115. ISBN: 978-3-319-23461-8.

Abstract: Experiments in high-energy astroparticle physics produce large amounts of data as continuous high-volume streams. Gaining insights from the observed data poses a number of challenges to data analysis at various steps in the analysis chain of the experiments. Machine learning methods have already cleaved their way selectively at some particular stages of the overall data mangling process.

Caron et al.: Analyzing {\gamma}-rays of the Galactic Center with Deep Learning
<div align="right">2017arXiv170806706C</div>

S. Caron et al. "Analyzing {\gamma}-rays of the Galactic Center with Deep Learning". In: *ArXiv e-prints* (Aug. 2017). arXiv: 1708.06706 [astro-ph.HE].

Abstract: We present a new method to interpret the γ-ray data of our inner Galaxy as measured by the Fermi Large Area Telescope (Fermi LAT). We train and test convolutional

neural networks with simulated Fermi-LAT images based on models tuned to real data. We use this method to investigate the origin of an excess emission of GeV γ-rays seen in previous studies. Interpretations of this excess include γ rays created by the annihilation of dark matter particles and γ rays originating from a collection of unresolved point sources, such as millisecond pulsars. Our new method allows precise measurements of the contribution and properties of an unresolved population of γ-ray point sources in the interstellar diffuse emission model.

Duarte et al.: Fast inference of deep neural networks in FPGAs for particle physics 2018arXiv180406913D

J. Duarte et al. "Fast inference of deep neural networks in FPGAs for particle physics". In: *ArXiv e-prints* (Apr. 2018). arXiv: 1804.06913 [physics.ins-det].

Abstract: Recent results at the Large Hadron Collider (LHC) have pointed to enhanced physics capabilities through the improvement of the real-time event processing techniques. Machine learning methods are ubiquitous and have proven to be very powerful in LHC physics, and particle physics as a whole. However, exploration of the use of such techniques in low-latency, low-power FPGA hardware has only just begun. FPGA-based trigger and data acquisition (DAQ) systems have extremely low, sub-microsecond latency requirements that are unique to particle physics. We present a case study for neural network inference in FPGAs focusing on a classifier for jet substructure which would enable, among many other physics scenarios, searches for new dark sector particles and novel measurements of the Higgs boson. While we focus on a specific example, the lessons are far-reaching. We develop a package based on High-Level Synthesis (HLS) called hls4ml to build machine learning models in FPGAs. The use of HLS increases accessibility across a broad user community and allows for a drastic decrease in firmware development time. We map out FPGA resource usage and latency versus neural network hyperparameters to identify the problems in particle physics that would benefit from performing neural network inference with FPGAs. For our example jet substructure model, we fit well within the available resources of modern FPGAs with a latency on the scale of 100 ns.

Feng et al.: The analysis of VERITAS muon images using convolutional neural networks feng_lin_2016

Qi Feng and Tony T. Y. Lin. "The analysis of VERITAS muon images using convolutional neural networks". In: *Proceedings of the International Astronomical Union* 12.S325 (2016), pp. 173–179. DOI: 10.1017/S1743921316012734.

Abstract: Imaging atmospheric Cherenkov telescopes (IACTs) are sensitive to rare gamma-ray photons, buried in the background of charged cosmic-ray (CR) particles, the flux of which is several orders of magnitude greater. The ability to separate gamma rays from CR particles is important, as it is directly related to the sensitivity of the instrument. This gamma-ray/CR-particle classification problem in IACT data analysis can be treated with the rapidly-advancing machine learning algorithms, which have the potential to outper-

form the traditional box-cut methods on image parameters. We present preliminary results of a precise classification of a small set of muon events using a convolutional neural networks model with the raw images as input features. We also show the possibility of using the convolutional neural networks model for regression problems, such as the radius and brightness measurement of muon events, which can be used to calibrate the throughput efficiency of IACTs.

D. George et al.: Glitch Classification and Clustering for LIGO with Deep Transfer Learning   2017arXiv171107468G

D. George, H. Shen, and E. A. Huerta. "Glitch Classification and Clustering for LIGO with Deep Transfer Learning". In: *ArXiv e-prints* (Nov. 2017). arXiv: `1711.07468` [`astro-ph.IM`].

Abstract: The detection of gravitational waves with LIGO and Virgo requires a detailed understanding of the response of these instruments in the presence of environmental and instrumental noise. Of particular interest is the study of anomalous non-Gaussian noise transients known as glitches, since their high occurrence rate in LIGO/Virgo data can obscure or even mimic true gravitational wave signals. Therefore, successfully identifying and excising glitches is of utmost importance to detect and characterize gravitational waves. In this article, we present the first application of Deep Learning combined with Transfer Learning for glitch classification, using real data from LIGO's first discovery campaign labeled by Gravity Spy, showing that knowledge from pre-trained models for real-world object recognition can be transferred for classifying spectrograms of glitches. We demonstrate that this method enables the optimal use of very deep convolutional neural networks for glitch classification given small unbalanced training datasets, significantly reduces the training time, and achieves state-of-the-art accuracy above 98.8%. Once trained via transfer learning, we show that the networks can be truncated and used as feature extractors for unsupervised clustering to automatically group together new classes of glitches and anomalies. This novel capability is of critical importance to identify and remove new types of glitches which will occur as the LIGO/Virgo detectors gradually attain design sensitivity.

D. George et al.: Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data   GEORGE201864

Daniel George and E.A. Huerta. "Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data". In: *Physics Letters B* 778 (2018), pp. 64–70. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/j.physletb.2017.12.053`. URL: `http://www.sciencedirect.com/science/article/pii/S0370269317310390`.

Abstract: The recent Nobel-prize-winning detections of gravitational waves from merging black holes and the subsequent detection of the collision of two neutron stars in coincidence with electromagnetic observations have inaugurated a new era of multimessenger astrophysics. To enhance the scope of this emergent field of science, we pioneered the use of deep learning with convolutional neural networks, that take time-series inputs, for rapid

detection and characterization of gravitational wave signals. This approach, Deep Filtering, was initially demonstrated using simulated LIGO noise. In this article, we present the extension of Deep Filtering using real data from LIGO, for both detection and parameter estimation of gravitational waves from binary black hole mergers using continuous data streams from multiple LIGO detectors. We demonstrate for the first time that machine learning can detect and estimate the true parameters of real events observed by LIGO. Our results show that Deep Filtering achieves similar sensitivities and lower errors compared to matched-filtering while being far more computationally efficient and more resilient to glitches, allowing real-time processing of weak time-series signals in non-stationary non-Gaussian noise with minimal resources, and also enables the detection of new classes of gravitational wave sources that may go unnoticed with existing detection algorithms. This unified framework for data analysis is ideally suited to enable coincident detection campaigns of gravitational waves and their multimessenger counterparts in real-time.

Holch et al.: Probing Convolutional Neural Networks for Event Reconstruction in Gamma-Ray Astronomy with Cherenkov Telescopes ICRC2017-Holch

T.L. Holch et al. "Probing Convolutional Neural Networks for Event Reconstruction in Gamma-Ray Astronomy with Cherenkov Telescopes". In: vol. 301. (ICRC2017)795. Proceedings of Science, 2017. URL: https://pos.sissa.it/301/795/.

Abstract: A dramatic progress in the field of computer vision has been made in recent years by applying deep learning techniques. State-of-the-art performance in image recognition is thereby reached with Convolutional Neural Networks (CNNs). CNNs are a powerful class of artificial neural networks, characterized by requiring fewer connections and free parameters than traditional neural networks and exploiting spatial symmetries in the input data. Moreover, CNNs have the ability to automatically extract general characteristic features from data sets and create abstract data representations which can perform very robust predictions. This suggests that experiments using Cherenkov telescopes could harness these powerful machine learning algorithms to improve the analysis of particle-induced air-showers, where the properties of primary shower particles are reconstructed from shower images recorded by the telescopes. In this work, we present initial results of a CNN-based analysis for background rejection and shower reconstruction, utilizing simulation data from the H.E.S.S. experiment. We concentrate on supervised training methods and outline the influence of image sampling on the performance of the CNN-model predictions.

IceCube Collaboration: Deep Learning in Physics exemplified by the Reconstruction of Muon-Neutrino Events in IceCube ICRC2017-Huennefeld

M. Huennefeld for the IceCube Collaboration. "Deep Learning in Physics exemplified by the Reconstruction of Muon-Neutrino Events in IceCube". In: vol. 301. (ICRC2017)1057. Proceedings of Science, 2017. URL: https://pos.sissa.it/301/1057/.

Abstract: Recent advances, especially in image recognition, have shown the capabilities of deep learning. Deep neural networks can be extremely powerful and their usage is compu-

tationally inexpensive once the networks are trained. While the main bottleneck for deep neural networks in the traditional domain of image classification is the lack of sufficient labeled data, this usually does not apply to physics where millions of Monte Carlo simulations exist. The IceCube Neutrino Observatory is a Cherenkov detector deep in the Antarctic ice where the reconstruction of muon-neutrino events is one of the key challenges. Due to limited computational resources and the high data rate, only simplified reconstructions limited to a small subset of data can be run on-site at the South Pole. However, in order to perform online analysis and to issue real-time alerts, a fast and powerful reconstruction is necessary. This paper demonstrates how deep learning techniques such as those used in image recognition can be applied to IceCube pulses in order to reconstruct muon-neutrino events. These methods can be generalized to other physics experiments.

Narayan et al.: Machine Learning-based Brokers for Real-time Classification of the LSST Alert Stream   2018arXiv180107323N

G. Narayan et al. "Machine Learning-based Brokers for Real-time Classification of the LSST Alert Stream". In: *ArXiv e-prints* (Jan. 2018). arXiv: `1801.07323` `[astro-ph.IM]`.

Abstract: The unprecedented volume and rate of transient events that will be discovered by the Large Synoptic Survey Telescope (LSST) demands that the astronomical community update its followup paradigm. Alert-brokers – automated software system to sift through, characterize, annotate and prioritize events for followup – will be critical tools for managing alert streams in the LSST era. The Arizona-NOAO Temporal Analysis and Response to Events System (ANTARES) is one such broker. In this work, we develop a machine learning pipeline to characterize and classify variable and transient sources only using the available multiband optical photometry. We describe three illustrative stages of the pipeline, serving the three goals of early, intermediate and retrospective classification of alerts. The first takes the form of variable vs transient categorization, the second, a multi-class typing of the combined variable and transient dataset, and the third, a purity-driven subtyping of a transient class. While several similar algorithms have proven themselves in simulations, we validate their performance on real observations for the first time. We quantitatively evaluate our pipeline on sparse, unevenly sampled, heteroskedastic data from various existing observational campaigns, and demonstrate very competitive classification performance. We describe our progress towards adapting the pipeline developed in this work into a real-time broker working on live alert streams from time-domain surveys.

Ryzhikov et al.: Domain adaptation with gradient reversal for MC/real data calibration   Ryzhikov2017

Domain adaptation with gradient reversal for MC/real data calibration. 2017. URL: `https://indico.cern.ch/event/567550/contributions/2629724/attachments/1513629/2515121/ACAT_PROCEEDINGS_2017.pdf`.

Abstract: It is quite common part of the data analysis in High Energy Physics to train a classifier for signal and background separation. In case the signal under investigation

is a rare process, the signal sample is simulated and background sample is taken from the real data. Such setting create an unnecessary bias: the classifier might learn not the characteristic of the signal but the characteristic of the imperfect simulation. So the challenge is to train the classifier in such way that it picks up signal/background difference and doesnt overfit to the simulation-specific features. The suggested approach is based on cross-domain adaptation technique using neural networks with gradient reversal. The network architecture is a dense multi-branch structure. One branch is responsible for the signal/background discrimination, the second branch helps to avoid the overfitting on the Monte-Carlo training dataset. The tests showed that this architecture is a robust mechanism for choosing trade-offs between discrimination power and overfitting. So the resulting networks successfully distinguishes the signal from the background, but does not distinguish simulated events from the real ones. The third network branch helps to reduce the correlation between the classifier predictions and reconstructed mass of the decay, thereby making such approach highly viable for wide variety of physics searches.

Shilon et al.: Application of Deep Learning methods to analysis of Imaging Atmospheric Cherenkov Telescopes data   2018arXiv180310698S

I. Shilon et al. "Application of Deep Learning methods to analysis of Imaging Atmospheric Cherenkov Telescopes data". In: *ArXiv e-prints* (Mar. 2018). arXiv: `1803.10698` `[astro-ph.IM]`.

Abstract: Ground based gamma-ray observations with Imaging Atmospheric Cherenkov Telescopes (IACTs) play a significant role in the discovery of very high energy (E > 100 GeV) gamma-ray emitters. The analysis of IACT data demands a highly efficient background rejection technique, as well as methods to accurately determine the energy of the recorded gamma-ray and the position of its source in the sky. We present results for background rejection and signal direction reconstruction from first studies of a novel data analysis scheme for IACT measurements. The new analysis is based on a set of Convolutional Neural Networks (CNNs) applied to images from the four H.E.S.S. phase-I telescopes. As the H.E.S.S. cameras pixels are arranged in a hexagonal array, we demonstrate two ways to use such image data to train CNNs: by resampling the images to a square grid and by applying modified convolution kernels that conserve the hexagonal grid properties. The networks were trained on sets of Monte-Carlo simulated events and tested on both simulations and measured data from the H.E.S.S. array. A comparison between the CNN analysis to current state-of-the-art algorithms reveals a clear improvement in background rejection performance. When applied to H.E.S.S. observation data, the CNN direction reconstruction performs at a similar level as traditional methods. These results serve as a proof-of-concept for the application of CNNs to the analysis of events recorded by IACTs.