

Deep Learning Transient Detection with VERITAS

Konstantin Pfrang^{a,*} on behalf of the VERITAS Collaboration

(a complete list of authors can be found at the end of the proceedings)

^a*Deutsches Elektronen Synchrotron DESY,
Platanenallee 6, 15738 Zeuthen, Germany
E-mail: konstantin.pfrang@desy.de*

Ground-based γ -ray observatories, such as the VERITAS array of imaging atmospheric Cherenkov telescopes, provide insight into very-high-energy (VHE, $E > 100$ GeV) astrophysical transient events. Examples include the evaporation of primordial black holes, gamma-ray bursts and flaring blazars. Identifying such events with a serendipitous location and time of occurrence is difficult. Thus, employing a robust search method becomes crucial. An implementation of a transient detection method based on deep-learning techniques for VERITAS will be presented. This data-driven approach significantly reduces the dependency on the characterization of the instrument response and the modelling of the expected transient signal. The response of the instrument is affected by various factors, such as the elevation of the source and the night sky background. The study of these effects allows enhancing the deep learning method with additional parameters to infer their influences on the data. This improves the performance and stability for a wide range of observational conditions. We illustrate our method for an historic flare of the blazar BL Lac that was detected by VERITAS in October 2016. We find a promising performance for the detection of such a flare in timescales of minutes that compares well with the VERITAS standard analysis.

*37th International Cosmic Ray Conference (ICRC 2021)
July 12th – 23rd, 2021
Online – Berlin, Germany*

*Presenter

1. Introduction

The detection of astrophysical transient events at very-high energies provides insight into different fundamental phenomena. These include previously detected transient phenomena such as gamma-ray bursts [1] or flaring blazars [2] and hypothesized transient phenomena such as the evaporation of primordial black holes [3]. The Very Energetic Radiation Imaging Telescope Array System (VERITAS) is an array of four imaging atmospheric Cherenkov telescopes (IACTs) located at the Fred Lawrence Whipple Observatory in southern Arizona (31 40 N, 110 57 W, 1.3 km a.s.l.) and is sensitive to sources of very-high-energy (VHE) photons above 100 GeV with a strength of 1% Crab Units (C.U.) in ~ 25 h [4]. VERITAS can detect VHE astrophysical transient sources within its field of view (FoV) of diameter of 3.5° [5].

IACTs detect an inevitable rate of background events, which pass the nominal selection cuts. The background includes γ -like cosmic-ray showers, electrons, positrons, and, to a lesser extent, diffuse γ rays. An astrophysical source may cause an enhanced event rate that over this background rate. Transient sources, cause a temporary increase of the observed event rates. These can appear at any location in the FoV and at any time. The observed rates change depending on the observation conditions. Thus, the changes in the instrument response need to be taken into account for a robust method to identify transients in the data. In this work, the approach proposed for IACTs in [6] is implemented for VERITAS. This deep-learning-based transient detection technique extracts information about the changes of the event rates from the data themselves. Being fully data-driven, it is insensitive to uncertainties on the modelling of the instrument and can operate across the entire FoV. It represents the first application of this method to real IACT data.

2. Deep Learning Transient Detection

Machine learning techniques are widely used in astronomy [7]. Among them, deep learning (DL) has proven its power in various astronomical applications (e.g., see [8]). It is based on deep and wide artificial neural networks (ANNs) that are able to represent complex models [9]. Recurrent Neural Networks (RNNs) are a type of ANNs that include cyclic connections that are especially suited to work with sequential data, such as time series. We use the framework presented in [6] that can be decomposed into a pair of encoder and decoder stages. As input, it takes a total of $\tau_{\text{RNN}} = \tau_{\text{enc}} + \tau_{\text{dec}}$ time steps. At each step, event counts N_{evt} in η energy bins are given as inputs to the network. The τ_{enc} steps of the encoder represent the background to a possible transient signal. Subsequently, the signal is searched for in the ensuing τ_{dec} decoder steps. Each time step of the RNN consists of a long short-term memory (LSTM) unit [10], comprised of a layer of 64 hidden units. The RNN is trained with sequences of time series which do not contain any signals (either in the encoder or in the decoder intervals). It thus learns to predict the expected background counts $B(\tau_{\text{dec}}, \eta)$. If a transient signal is present as part of the decoder interval, the total observed counts $S(\tau_{\text{dec}}, \eta)$ are potentially increased. In this work, we use a different test statistic (TS) than the one used for [6], who assumed that the inputs to the network followed Poissonian statistics. We instead define our TS as

$$\text{TS}(\eta) = \frac{S(\tau_{\text{dec}}, \eta) - B(\tau_{\text{dec}}, \eta)}{\sqrt{|B(\tau_{\text{dec}}, \eta)| + 1}}. \quad (1)$$

The one in the denominator is added to avoid numerical division by zero. The TS is independently computed for the event counts in each energy bin. As the training data consists exclusively of background sequences, it can be used to map TS values into a p -value for detection. We define the combined TS as the average of $|\log(p)|$ over all features and map it to a combined p -value itself.

3. Data Selection and Preparation

This work utilizes VERITAS data recorded from 2012 to 2020. To learn the changes in the background rates, a dataset that does not include uncontrolled systematic effects is required. Possible systematics include hardware issues or the effects of bad weather, e.g., by clouds in the field of view. As these result in a drop of the event rates, they potentially degrade the reliability of a search for transients.

For the generation of the training dataset, we apply a selection on data-taking runs, based on existing data quality monitoring results and basic filters. We consider data that was taken while operating the full four-telescope array at zenith angles below 45° with nominal high voltage settings and without recorded hardware issues. Manual data quality checks for a large dataset are time-consuming and a possible source of inconsistency. To avoid this, we deploy an automatic pipeline, that assesses the data quality after the initial selection. We investigate time series data, such as the array trigger (L3) rate and temperature of the sky for systematic effects. Among these are spikes or drops in the trigger rate e.g., due to a short flash of light in the camera, clouds in the FoV, and changes correlated with the amount of night sky background (NSB). When such substantial systematics are detected, time cuts are applied and the corresponding interval in the data is discarded. We validate the remaining intervals for consistency with a constant trigger rate to indicate a stable performance of VERITAS, where a final time cut may also be applied. Following this procedure, we obtain a training dataset with an integrated $\sim 2,730$ hours of observations.

The data are processed using one of the standard VERITAS analysis packages [11]. We start from a list of events following the standard event reconstruction, e.g., having reconstructed energies and directions of arrival. For each run, we define our ROIs as non-overlapping circles, up to a distance of 1.5° from the center of the FOV. See Figure 1. We choose a radius for the ROIs of 0.25° . This conservatively covers the point spread function of VERITAS (uncertainty on event localisation), over our chosen event parameter range. Bright stars and known VHE γ -ray sources in the FoV are masked with exclusion regions. Events coming from these directions are not considered for the analysis. The observations are binned into time steps of 30 s intervals. Time cuts that are determined by the automatic quality assessment are applied. If more than 10% of a ROI or a time step are masked, they are removed completely. Otherwise we correct the measured events for the expected missing fraction, assuming uniform distributions within each ROI. In each of the ROIs and time steps, the count of events is calculated in three energy bins, 100 GeV to 330 GeV, 330 GeV to 1 TeV, and 1 TeV to 100 TeV.

For γ -hadron separation (classification of events as non- γ -ray background), boosted decision trees (BDTs) are utilized [12]. These assign a score between -1 and 1 to each event, which indicates how " γ -like" it appears. Depending on the zenith and energy bin, events with $\text{BDT} \geq 0.34$ to $\text{BDT} \geq 0.71$ are selected in the standard VERITAS analysis. In this work, we deploy two different counting schemes to obtain the number of events N_{evt} per bin. In the first approach, we do not

distance accounts for the difference in the radial acceptance. – Finally, the multiplicity in the different energy bins contributes to identifying differences in the energy threshold that affects the rate of low-energy events.

Table 1: Auxiliary parameters and meta bins for prediction of the background counts.

<i>Parameter</i>	<i>Description</i>	<i>Used</i>	<i>Meta bins</i>
l3_mean	Mean L3 trigger rate during time step	No	–
nsb_level	Mean charge in camera indicating NSB	No	–
azimuth	Azimuth of pointing position	Yes	North, South
$\sec(\theta)$	secant of the pointing zenith angle	Yes	[1.0, 1.08), [1.08, 1.16), [1.16, 1.24), [1.24, 1.4)
ref_time	Time after August 1, 2012 in years	Yes	2012-13, 2013-14, 2014-20
offset	Distance of ROI to camera center	Yes	6 bins, maximum 10% difference in radial acceptance
multiplicity(η)	Average number of images at each time step and energy bin	Yes	–

As the expected background counts change, the interpretation of the TS can be adjusted. We introduce meta bins to split the observations into chunks in which the observed background rates are mostly stable. The study to determine the auxiliary parameters revealed that the most critical parameters are the azimuth, $\sec(\theta)$, ref_time, and the offset. In total, we define 144 bins that are summarized in Table 1.

5. Detection of BL Lac Flare

In its low state, the blazar BL Lac is not detectable in VERITAS over the timescales investigated in this work (few minutes). However, in the past it proved to be a strong transient source, exhibiting flaring periods. We investigate the BL Lac flare which occurred during October 2016 that reached a flux of ~ 1.8 C.U. above 200 GeV in the 4-minute-binned VERITAS light curve [2]. In this work, we consider the 30 minute run that corresponds to the highest flux in the 30-minute-binned light curve (see Figure 1 of [2]). We investigate the potential to detect such a flare with the DL transient detection method.

The decoder steps, τ_{dec} , in which the signal is searched are filled with data obtained during the flaring period. The DL transient method requires inputs for the encoder steps, τ_{enc} , that represent the background event rates. As the source is in the flaring state, we generate the encoder counts by sampling from shuffled BL Lac observations during a low state, taken under similar observing conditions. Generally, the padding for the encoder steps can also be extracted directly from the training dataset, within the same meta bins. The light curves for both counting schemes are shown in Figure 2.

We iteratively increase the number of decoder steps from 1 to 60, investigating the change in detection significance as a function of the observation time t_{obs} of the flare. The results are shown as the blue curves in Figure 3. In this study, the padding of the encoder samples is repeated 1000

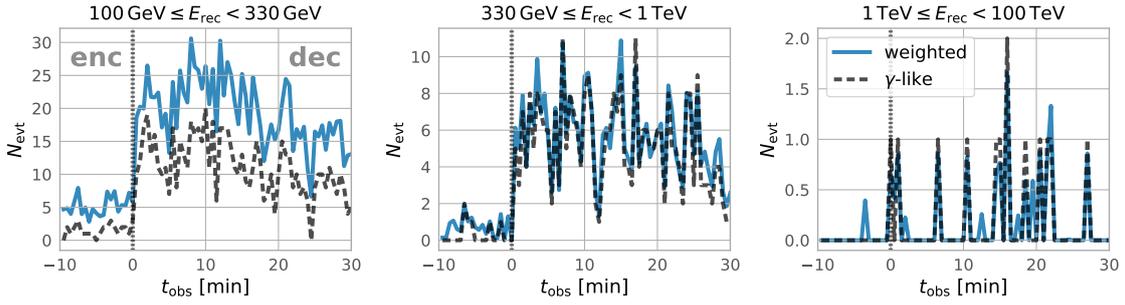


Figure 2: Number of events N_{evt} as function of the time t_{obs} in the generated light curve of BL Lac in three energy bins. The transition between the encoder and the decoder interval is at $t_{\text{obs}} = 0$. The solid blue curves represent the counts obtained from the weighted sum after the loose BDT cut. The dashed black lines are the counts of events after the γ -hadron separation. The encoder steps are sampled from a shuffled low state dataset and the decoder shows the counts during the flare.

times in order to derive the corresponding uncertainties (95% containment). On the left graph, we investigate the individual contributions of the 6 input features to the combined significance. One may observe that the significance for the two different counting metrics are comparable. Overall the γ -like counts perform slightly better on longer timescales. This reflects the higher signal to noise ratio, as a larger fraction of the background is discarded. For short timescales and at highest energies, the weighted counting metric is slightly more informative for the network, given that this regime is dominated by statistical uncertainties.

The TS is calculated and mapped to p-values for each of the six input counts independently. This allows our method to also perform well in cases where only a subset of the features becomes particularly important. For example a very hard source spectrum could lead to a large enhancement of the γ -ray rate in the highest energy bin. As a very low background rate is expected in this case, the related significance would be comparably higher. On the right side of Figure 3, we compare the combined significance against the results obtained for $E > 100$ GeV with the VERITAS standard analysis that uses equation 17 of [14]. We use the *reflected region* background method [15]. For this, *off regions* with the same size as the target region are defined on a circle around the camera center. The background is contemporaneously determined from the number of γ -like events arriving from these regions. The background estimation for the RNN is based on the encoder steps of one ROI, in an area corresponding to a single off region. The significance obtained with our method is compatible with the results of the standard analysis throughout most of the run. For short timescales, the reflected region method is limited by low background statistics. On the other hand, the DL method predicts the counts based on previous observations, and so is more robust. Thus, our approach has the potential to outperform the standard method on short time scales. This can be critical when dealing with transient signals that might have rapid variability or short timescales.

We also show for illustration results for the standard VERITAS analysis using six off regions. A higher number of off regions lead to better estimation of the background, and correspondingly to higher detection significance. In general, the definition of off regions in the FoV can be challenging for geometrical reasons or the presence of γ -ray sources. Thus, this method is often used for detailed offline analyses instead for blind searches across the entire FoV. In future studies, we

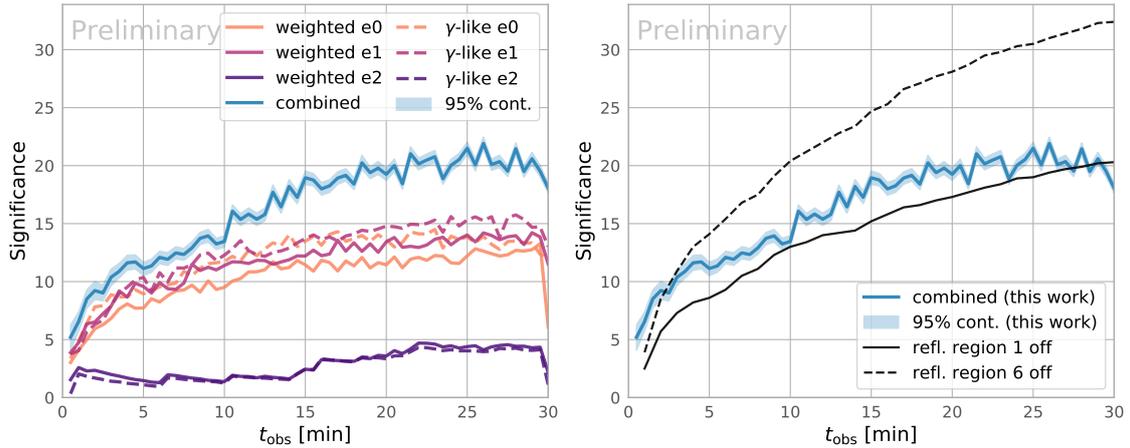


Figure 3: Significance of detection as function of the observation time of the flare t_{obs} . The calculation is repeated 1000 times while resampling the encoder steps. We show the medium and the 95% containment interval in blue. The left graph shows the individual contribution to the significance of all 6 parameters. Solid lines represent the weighted counts after the loose BDT cuts and the dashed are counts of γ -like events. The individual colors show the three energy bins. The black reference curves in the right are derived from the VERITAS standard analysis above 100 GeV using the reflected region background with 1 and 6 off regions respectively.

will investigate alternative training schemes for our RNN, which might take advantage of larger background estimation areas. However, we also note that on short timescales, even given the addition of off regions, the DL approach performs slightly better in comparison.

6. Summary

In this study we present the first implementation of a deep learning based method to detect transient sources with VERITAS. We developed a pipeline to perform an automatic data quality assessment and convert the event lists of the standard analysis to the inputs. The required auxiliary parameters and meta bins were determined by studying their effects on the observed background rates. As part of the anomaly detection approach presented in this work, the network learns to predict the expected background rates. A transient signal can be detected as a divergence from these predictions. We present a preliminary choice of selected inputs to such a network.

We illustrate our methodology on an historical flare of the blazar BL Lac, which was observed during October 2016. The generated time series correspond to a scenario of a possible follow-up observation. Overall the results are compatible with the significance achieved by traditional detection methods, considering a single reflected background region. The results for short timescales are promising, which can be critical for the fast detection of transient signals.

Acknowledgements

This research is supported by grants from the U.S. Department of Energy Office of Science, the U.S. National Science Foundation and the Smithsonian Institution, by NSERC in Canada, and by

the Helmholtz Association in Germany. This research used resources provided by the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science, and resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. We acknowledge the excellent work of the technical support staff at the Fred Lawrence Whipple Observatory and at the collaborating institutions in the construction and operation of the instrument. K. Pfrang acknowledges the support of the Young Investigators Program of the Helmholtz Association.

References

- [1] H. Abdalla, R. Adam, F. Aharonian, et al. A very-high-energy component deep in the γ -ray burst afterglow. *Nature*, 575:464–467, 2019.
- [2] A. U. Abeysekara, et al. Multiwavelength observations of the blazar BL lacertae: A new fast TeV gamma-ray flare. *The Astrophysical Journal*, 856(2):95, Mar 2018.
- [3] B. J. Carr et al. New cosmological constraints on primordial black holes. *Phys. Rev. D*, 81:104019, May 2010.
- [4] J. Holder. VERITAS: Status and highlights. *ICRC2011*, 11, 2011.
- [5] J. Holder et al. The first VERITAS telescope. *Astroparticle Physics*, 25(6):391–401, 2006.
- [6] I. Sadeh. Data-driven Detection of Multimessenger Transients. *Astroph. J.*, 894(2):L25, 2020.
- [7] N. M. Ball and R. J. Brunner. Data mining and machine learning in astronomy. *International Journal of Modern Physics D*, 19(07):1049–1106, 2010.
- [8] G. Carleo et al. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91:045002, 2019.
- [9] Y. LeCun, Y. Bengio, and G Hinton. Machine learning and the physical sciences. *Nature*, 521:426–444, 2015.
- [10] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9:1735–1780, 1997.
- [11] G. Maier and J. Holder. Eventdisplay: An Analysis and Reconstruction Package for Ground-based Gamma-ray Astronomy. *ICRC2017*, 301, 2017.
- [12] M. Krause, E. Pueschel, and G. Maier. Improved γ /hadron separation for the detection of faint γ -ray sources using boosted decision trees. *Astroparticle Physics*, 89:1–9, 2017.
- [13] M. N. Rosillo et al. (VERITAS collaboration). The throughput calibration of the VERITAS telescopes (in preparation). 2021.
- [14] T. P. Li and Y. Q. Ma. Analysis methods for results in gamma-ray astronomy. *Ap. J.*, 272:317–324, September 1983.
- [15] Berge, D., Funk, S., and Hinton, J. Background modelling in very-high-energy astronomy. *A&A*, 466(3):1219–1229, 2007.

Full Authors List: VERITAS Collaboration

C. B. Adams¹, A. Archer², W. Benbow³, A. Brill¹, J. H. Buckley⁴, M. Capasso⁵, J. L. Christiansen⁶, A. J. Chromey⁷, M. Errando⁴, A. Falcone⁸, K. A. Farrell⁹, Q. Feng⁵, G. M. Foote¹⁰, L. Fortson¹¹, A. Furniss¹², A. Gent¹³, G. H. Gillanders¹⁴, C. Giuri¹⁵, O. Gueta¹⁵, D. Hanna¹⁶, O. Hervet¹⁷, J. Holder¹⁰, B. Hona¹⁸, T. B. Humensky¹, W. Jin¹⁹, P. Kaaret²⁰, M. Kertzman², T. K. Kleiner¹⁵, S. Kumar¹⁶, M. J. Lang¹⁴, M. Lundy¹⁶, G. Maier¹⁵, C. E McGrath⁹, P. Moriarty¹⁴, R. Mukherjee⁵, D. Nieto²¹, M. Nievas-Rosillo¹⁵, S. O'Brien¹⁶, R. A. Ong²², A. N. Otte¹³, S. R. Patel¹⁵, S. Patel²⁰, K. Pfrang¹⁵, M. Pohl^{23,15}, R. R. Prado¹⁵, E. Pueschel¹⁵, J. Quinn⁹, K. Ragan¹⁶, P. T. Reynolds²⁴, D. Ribeiro¹, E. Roache³, J. L. Ryan²², I. Sadeh¹⁵, M. Santander¹⁹, G. H. Sembroski²⁵, R. Shang²², D. Tak¹⁵, V. V. Vassiliev²², A. Weinstein⁷, D. A. Williams¹⁷, and T. J. Williamson¹⁰

¹Physics Department, Columbia University, New York, NY 10027, USA ²Department of Physics and Astronomy, DePauw University, Greencastle, IN 46135-0037, USA ³Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA 02138, USA ⁴Department of Physics, Washington University, St. Louis, MO 63130, USA ⁵Department of Physics and Astronomy, Barnard College, Columbia University, NY 10027, USA ⁶Physics Department, California Polytechnic State University, San Luis Obispo, CA 94307, USA ⁷Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA ⁸Department of Astronomy and Astrophysics, 525 Davey Lab, Pennsylvania State University, University Park, PA 16802, USA ⁹School of Physics, University College Dublin, Belfield, Dublin 4, Ireland ¹⁰Department of Physics and Astronomy and the Bartol Research Institute, University of Delaware, Newark, DE 19716, USA ¹¹School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA ¹²Department of Physics, California State University - East Bay, Hayward, CA 94542, USA ¹³School of Physics and Center for Relativistic Astrophysics, Georgia Institute of Technology, 837 State Street NW, Atlanta, GA 30332-0430 ¹⁴School of Physics, National University of Ireland Galway, University Road, Galway, Ireland ¹⁵DESY, Platanenallee 6, 15738 Zeuthen, Germany ¹⁶Physics Department, McGill University, Montreal, QC H3A 2T8, Canada ¹⁷Santa Cruz Institute for Particle Physics and Department of Physics, University of California, Santa Cruz, CA 95064, USA ¹⁸Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA ¹⁹Department of Physics and Astronomy, University of Alabama, Tuscaloosa, AL 35487, USA ²⁰Department of Physics and Astronomy, University of Iowa, Van Allen Hall, Iowa City, IA 52242, USA ²¹Institute of Particle and Cosmos Physics, Universidad Complutense de Madrid, 28040 Madrid, Spain ²²Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA ²³Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam-Golm, Germany ²⁴Department of Physical Sciences, Munster Technological University, Bishopstown, Cork, T12 P928, Ireland ²⁵Department of Physics and Astronomy, Purdue University, West Lafayette, IN 47907, USA