Astroparticle Physics 97 (2018) 46–53

Contents lists available at ScienceDirect

Astroparticle Physics

journal homepage: www.elsevier.com/locate/astropartphys

A deep learning-based reconstruction of cosmic ray-induced air showers

M. Erdmann*, J. Glombitza, D. Walz

Physics Institute 3A, RWTH Aachen University, Aachen D-52056, Germany

ARTICLE INFO

Article history:
Received 2 August 2017
Revised 7 September 2017
Accepted 18 October 2017
Available online 20 October 2017

Keywords:
Cosmic rays
UHECR
Air showers
Deep learning

ABSTRACT

We describe a method of reconstructing air showers induced by cosmic rays using deep learning techniques. We simulate an observatory consisting of ground-based particle detectors with fixed locations on a regular grid. The detector's responses to traversing shower particles are signal amplitudes as a function of time, which provide information on transverse and longitudinal shower properties. In order to take advantage of convolutional network techniques specialized in local pattern recognition, we convert all information to the image-like grid of the detectors. In this way, multiple features, such as arrival times of the first particles and optimized characterizations of time traces, are processed by the network. The reconstruction quality of the cosmic ray arrival direction turns out to be competitive with an analytic reconstruction algorithm. The reconstructed shower direction, energy and shower depth show the expected improvement in resolution for higher cosmic ray energy.

1. Introduction

The successes of the deep learning ansatz in handwriting recognition [1,2], speech recognition [3], and competitions with humans regarding image identification and playing games [4–6] motivates its application in various areas of fundamental research. For an overview of deep learning techniques see [7].

In physics research, tasks like reconstructing particle kinematics or signal separation from background processes often require algorithms that take multiple observables into account in order to obtain a few parameters of interest. Machine learning techniques are frequently used for such complex problems, with boosted decision trees and shallow neural networks being among the most popular methods.

Concepts of deep neural networks have recently been investigated for challenges in astroparticle and particle physics. Various network designs have been successfully applied in simulation studies, e.g., to reconstruct the neutrino flavor in neutrino nucleus interactions [8], to extract a new exotic particle or Higgs boson signal from a background dominated data sample [9–11], to identify the underlying parton flavor of a jet or measure jet substructure [12,13], or to assign jets to the underlying partons in top quark-associated Higgs boson production [14].

In this work we investigate deep learning methods for reconstructing cosmic ray properties from simulated air showers. As our observatory of the air showers we use ground-based particle detectors located on a hexagonal grid, each delivering charge measurements as a function of time.

Reconstruction of the cosmic ray properties from air showers includes arrival direction and energy. In addition, the atmospheric depth of the shower maximum is of interest to obtain information on the cosmic ray mass. All this information is commonly extracted from the data using algorithms based on physics arguments which were developed by astroparticle physicists.

As an alternative method, we exploit deep learning techniques to reconstruct the above-mentioned properties of cosmic rays from simulated data. Our aim is to investigate the abilities of the network to learn about various aspects encoded in the data, and to evaluate the reconstruction quality of the cosmic ray arrival direction, energy, and shower maximum.

We designed the network to study the following aspects separately. One aspect is the spatial distribution of signals on the ground. Further information is contained in the time of the first signals arriving at different detector stations on the ground. A third aspect is the time trace of the detector signals which encodes arrival times of different particles from different phases of the shower development.

This work is structured as follows. First we introduce a parametrized simulation of air showers. We then explain the details of the network architecture and discuss how the data are in-

* Corresponding author.
E-mail address: erdmann@physik.rwth-aachen.de (M. Erdmann).

https://doi.org/10.1016/j.astropartphys.2017.10.006

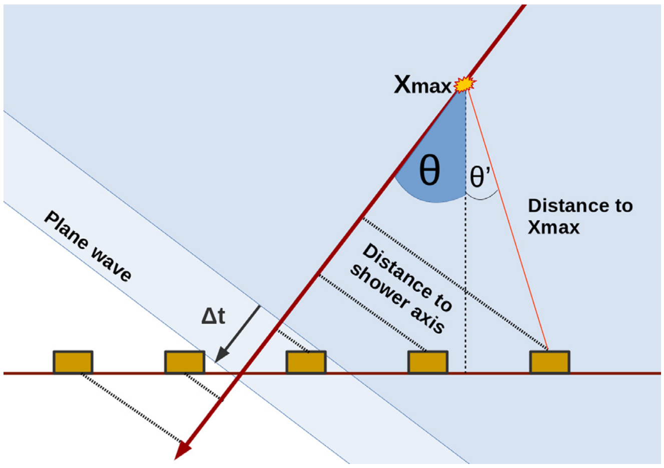

Fig. 1. Parametrized air shower simulation with zenith angle θ and reference point $X_{\max}$ to determine the detector distances to the shower axis and to estimate atmospheric absorption effects. A planar shower front is used to determine the arrival times of the first particles at the detectors.

put to the network. We evaluate the quality of air shower reconstructions obtained with the network using various aspects of the data before presenting our conclusions.

2. Parametrized simulation of air showers

For our investigations we developed a parametrized air shower simulation code, and a corresponding detector simulation inspired by [15,16]. The simulation code is published in [17]. This program enables the efficient generation of a large number of simulated events for network training and evaluation on one hand, and direct control of the information contained in the data on the other hand. On output the simulation delivers the same information as obtained by a ground-based particle detector such that there are no limitations in extending our studies to detailed air shower simulations.

The results of the air shower simulations are parametrized directly in terms of the signal distributions obtained in the detectors. The detectors are placed on a hexagonal grid with a spacing of 1500 m and are located at a height of 1400 m above sea level, motivated by the Pierre Auger Observatory [18].

We consider only the electromagnetic part and the muon part of the air shower, both of which depend on the arrival directions, energies, and nuclear masses of the cosmic rays.

The energy E of the cosmic rays is randomly selected between 3 and 100 EeV following a power law E^{-1}. The mass composition A of the cosmic rays consists of H, He, N, Fe with equal fractions. The arrival directions (θ, ϕ) are chosen following an isotropic distribution with zenith angles of up to $\theta = 60°$. Finally, the impact point of the shower on the surface, the shower core, is randomly sampled around the central detector.

For the chosen values of the energy and mass (E, A), two decisions are taken which influence the detector signals. The first decision is the spatial reference point for calculating distances of the detectors to the shower, see Fig. 1. The distribution of the maximum of the atmospheric shower depth $X_{\max}$ can be well approximated by a Gumbel distribution $G(E, A)$ [19]. The randomly selected $X_{\max}$ value from $G(E, A)$, together with the arrival direction (θ, ϕ) and the shower core, defines the reference point for all subsequent geometry calculations. Starting here, the movement of the shower front is approximated in terms of a plane wave.

The second decision to depend on the chosen (E, A) values is the relative energy distribution between the electromagnetic and the muonic components of the shower. For proton primaries, we set the electromagnetic energy to 70% of the cosmic ray energy and the muon part to 30%, respectively. For all other nuclei, the relative energy contained in the muonic component is up-scaled by a factor $A^{0.15}$ [20]. The signal distributions in the detectors of both the electromagnetic and the muonic components are calculated separately and are finally superimposed.

Each detector is supposed to have its own clock recording a universal time t_0 when the trigger starts the electronics to record particle signals as a function of time t. The time t_0 is calculated according to the movement of the shower front.

The time trace $S(t)$ of particle signals arriving at the detector contains 80 intervals of 25 ns size. The shape of the signal distributions is approximated by a log-normal distribution for both the electromagnetic and the muon signal ($j = \mu, em$), following [16]:

$$F_j(t) = \frac{1}{\sqrt{2\pi\sigma^2}\, x_j} \exp\left(-\frac{(\ln(x_j) - \tau_j)^2}{2\sigma^2}\right) \qquad (1)$$

with $x_j = (t - \Delta t_j)/t'$. Here t' is a reference time to cancel time units, Δt_j a time offset specified below, τ_j the location parameter of the log-normal distribution and $\sigma = 0.7$ the shape parameter. To simulate the effect of a lateral distribution function of the shower, the location parameter

$$\tau_j = \ln\left[a_j + b_j \left(\frac{r}{r_\circ}\right)^c \left(1 - d_j\, \frac{\Delta X}{X_\circ}\right)\right] \qquad (2)$$

depends on the transverse distance r of the detector to the shower axis. To also include the absorption of shower particles in the atmosphere, the difference ΔX between the atmospheric depth of the detector and the above-mentioned reference point of the shower at $X_{\max}$ is calculated. a_j contains a global offset, and b_j and d_j are weights for the lateral distance and the absorption effects, which are different for the muonic and the electromagnetic components. The choice of parameters has been adapted to approximate distributions obtained from full shower simulations [16] ($a_\mu = 80,\ a_{em} = 50,\ b_\mu = 140,\ b_{em} = 200\ ,\ c = 1.4,\ d_\mu = 0.2,\ d_{em} = 0.1,\ r_\circ = 750\,\mathrm{m},\ X_\circ = 1000\,\mathrm{g/cm^2}$). For the electromagnetic component a time offset is added, reading

$$\frac{\Delta t_{em}}{t'} = \frac{3}{2}\left(\exp(\tau_{em}) - \exp(\tau_\mu)\right) \qquad (3)$$

with respect to the muonic component, $\Delta t_\mu / t' = 0$.

The total energy deposit in the detector, or total signal strength S'_0 respectively, respects the same effects of the lateral distribution function and the atmospheric absorption as presented in (2):

$$S'_{j,0} = S_{j,0} \left(\frac{r}{r'}\right)^{\alpha_j} \left(\frac{\Delta X}{X'}\right)^{\beta_j}. \qquad (4)$$

Here the parameters are $\alpha_\mu = -4.7,\ \alpha_{em} = -6.1,\ \beta_\mu = 0.1,\ \beta_{em} = 0.4,\ r' = 1000\,\mathrm{m},\ X' = 100\,\mathrm{g/cm^2}$.

For each detector i, the signal distribution $S_i(t)$ as a function of time is obtained by a Monte Carlo method. Random values are chosen according to (1) where the number of drawn values is proportional to the signal strength $S'_{j,0}$ (see Eq. (4)). To prevent diverging signals a maximum signal is set with respect to the shower core.

Finally, effects of noise contributions are added to the signal distributions. In each interval of the time trace a uniformly distributed noise between $0 - 5\%$ of the signal and a Gaussian background noise are added to the interval ($\mu = 1.2,\ \sigma = 0.6$).

Also the trigger times t_0 of each detector are subjected to noise effects expressed in terms of Gaussian distributions where the width varies between $0.06\mu s$ and $0.32\mu s$ depending on the total signal ($S'_0 = S'_{em,0} + S'_{\mu,0}$) in the detector.

Fig. 2 shows example distributions of the simulated time traces for a single event in a detector near the shower core (Fig. 2a), and in a detector at some distance from the shower core (Fig. 2b).