MACHINE LEARNING
Science and Technology

Neural networks for separation of cosmic gamma rays and hadronic cosmic rays in air shower observation with a large area surface detector array

Sousuke Okukawa[1,*] , Kazuyuki Hara[2], Kinya Hibino[3] , Yusaku Katayose[1] , Kazumasa Kawata[4], Munehiro Ohnishi[4], Takashi Sako[4] , Takashi K Sako[4], Makio Shibata[1], Atsushi Shiomi[2] and Masato Takita[4]

[1] Faculty of Engineering, Yokohama National University, Yokohama 240-8501, Japan
[2] College of Industrial Technology, Nihon University, Narashino 275-8575, Japan
[3] Faculty of Engineering, Kanagawa University, Yokohama 221-8686, Japan
[4] Institute for Cosmic Ray Research, University of Tokyo, Kashiwa 277-8582, Japan
* Author to whom any correspondence should be addressed.

E-mail: okukawa-sousuke-dt@ynu.jp

Abstract

The Tibet ASγ experiment has been observing cosmic gamma rays and cosmic rays in the energy range from teraelectron volts to several tens of petaelectron volts with a surface detector array since 1990. The derivation of cosmic gamma-ray flux is made by finding the excess distribution of the arrival direction of air showers above background cosmic rays. In 2014, the underground water Cherenkov muon detector (MD) was added to separate cosmic gamma rays from the background on the basis of the muon-less feature of the air showers of gamma-ray origin; hybrid observations using these two detectors were started at this time. In the present study, we developed methods to separate gamma-ray-induced air showers and hadronic cosmic-ray-induced ones using the measured particle number density distribution to improve the sensitivity of cosmic gamma-ray measurements using the Tibet air shower array data alone before the installation of the MD. We tested two approaches based on neural networks. The first method used feature values representing the lateral spread of the secondary particles, and the second method used the shower image data. To compare the separation performance of each method, we analyzed Monte Carlo air shower events in the vertically incident direction with mono-initial-energy gamma rays and protons. When discriminated by a single feature, the feature with the highest separation performance has an area under the curve (AUC) value of 0.701 for a gamma-ray energy of 10 TeV and 0.808 for 100 TeV. A separation method with a multilayer perceptron (MLP) based on multiple features has AUC values of 0.761 for a gamma-ray energy of 10 TeV and 0.854 for 100 TeV, which represents an improvement of approximately 5% in the AUC value compared with the single-feature case. We also found that the feature values that effectively contribute to the separation vary depending on the energy. A separation method with a convolutional neural network (CNN) using the shower image data has AUC values of 0.781 for a gamma-ray energy of 10 TeV and 0.901 for 100 TeV, which are approximately 5% higher than those of the MLP method. We applied the CNN separation method to Monte Carlo gamma-ray and cosmic-ray events from the Crab Nebula in the energy range 10–100 TeV. The AUC values range from 0.753 to 0.879, and the significance of the observed gamma-ray excess is improved by 1.3 to 1.8 times compared with the case without the separation procedure.

1. Introduction

The acceleration mechanism of charged particles in astrophysical objects and the generation mechanism of electromagnetic waves are critical topics for understanding the physical processes within those objects. In addition, electromagnetic waves carry information on astronomical phenomena. Astronomical observations using electromagnetic waves have been carried out using multi-wavelength observations of radiation ranging from radio waves to gamma rays with energies as high as several tens of teraelectron volts (TeV). In recent years, there has been significant progress in the development of techniques for observing cosmic gamma rays in the sub-petaelectron volt region, and these techniques have become critical for studying the origin of cosmic rays and astrophysical phenomena [1–3]. The Tibet ASγ group has revealed the existence of gamma-ray emissions exceeding 100 TeV from the Crab Nebula for the first time [1]. One of the difficulties in measuring such high-energy gamma rays is the low arrival frequency. Therefore, a large-area detector is essential for carrying out high statistics measurements. However, high-energy cosmic gamma rays ('primary gamma rays') reaching the earth create many secondary particles by colliding with atmospheric nuclei—a phenomenon known as an 'air shower.' The number of particles in an air shower, starting above the atmosphere, increases with decreasing altitude and reaches a maximum at a certain altitude. The number of particles then begins to decrease because of the energy loss of each particle. Consequently, several experimental groups, including the Tibet ASγ group, have been conducting measurements with large-area detectors installed on the ground ('ground-based detectors') at altitudes greater than 4000 m above sea level, which is close to the maximum development altitude of air showers, to enhance the energy and directional resolution of the measurements. Another difficulty with observing primary gamma rays is the enormous amount of background cosmic rays such as protons and helium. These cosmic rays also produce air showers. Observing high-energy gamma-ray sources also necessitates capturing the faint gamma-ray-induced air showers buried within the vast background of hadronic cosmic-ray-induced ones.

Cosmic gamma rays generate a pure shower of electromagnetic particles (i.e. secondary gamma rays and secondary electrons) by electromagnetic interaction; this phenomenon is known as an 'electromagnetic shower.' However, nuclei are groups of composite particles called hadrons bound together by strong interactions. In the initial stage of the air shower of a hadronic cosmic ray (a 'cosmic-ray-induced air shower'), the strong interactions increase the chain of production of mesons (a type of composite particle bound by strong interactions) such as π mesons and kaons. Among these, the neutral π meson (π^0) decays into two gamma rays, creating multiple electromagnetic showers [4].

In ground-based experiments with surface-type air shower arrays such as the TibetASγ group, the separation of gamma-ray-induced air showers (hereafter called gamma-showers) and hadronic cosmic-ray-induced ones (hereafter called hadronic-showers or proton-showers in the case of proton origin) is primarily performed through two methods. One of these methods involves using the muon intensity in the air shower. The Tibet ASγ group installed a large underground water-Cherenkov-type muon detector (MD) beneath the surface particle detector array in 2014, enabling the separation of gamma-showers and hadronic-showers on the basis of the muon abundance in the air shower [5, 6]. The second method involves examining the density distribution of secondary particles in an air shower detected by surface detectors. Gamma-showers are almost pure electromagnetic showers, leading to simple structures in the lateral spread distribution of secondary particle density reaching the ground. By contrast, hadronic-showers consist of multiple sub-showers originating from gamma rays produced by π^0 mesons created through nuclear interactions in the upper atmosphere, leading to complicated lateral structures. Therefore, the gamma-showers and the hadronic-showers can be separated by measuring the differences in the particle density distribution. For example, HAWC employed a water Cherenkov air shower array with a depth of 4 m and developed a separation method using a feature value called 'compactness,' which represents the high-particle-density region that arises from muons and sub-showers and is formed away from the core of the air shower [7]. The Astrophysical Radiation with Ground-based Observatory at YangBaJing (ARGO-YBJ) experiment, by comparison, used a carpet-type array of resistive plate chambers and applied a feed-forward neural network to multiple feature values such as the number of hit detectors for the separation [8]. The Tibet ASγ project installed an MD in 2014; however, gamma-ray observations were conducted using the Tibet air shower array alone before 2014. A reanalysis of data accumulated over a long period without MD for transient phenomena such as flares [9, 10] using a new separation method would be informative.

In recent years, machine learning algorithms, especially the method known as deep learning, which is a multilayered neural network, have made great strides [11]. The applications of machine learning for image identification are vast and diverse and include object detection, face recognition, medical-image analysis, scene recognition, and optical character recognition. The high effectiveness of this technology has been demonstrated in a wide range of tasks. Neural network-based analysis methods have been developed for air shower observations using ground-based detectors, such as the Pierre Auger experiment and the Telescope

Array experiment, and have demonstrated superior performance compared with conventional analysis methods [12, 13].

We developed a neural network-based method to separate gamma and hadron showers (γ/CR-separation). We also attempted to apply convolutional neural networks (CNN), which were developed for image identification, to image data created from physical quantities such as charged particle density distributions measured in air shower arrays. This paper summarizes the detailed methodology and physical interpretations of the results of our previous work [11] and adds new results. In section 2, we briefly describe the configuration of the Tibet air shower array and its method of observing air showers. In section 3, we applied the two separation methods to particle density data collected using the Tibet air shower array. To compare their separation performance, we analyzed Monte Carlo data of vertically incident gamma-showers and proton-showers with mono initial energies (γ/p-separation). One method used specific feature values defined with the particle number density data. First, we tested only a single feature value without a neural network. We then tested a multilayer perceptron (MLP) neural network with multiple feature values. The other method involved image data converted from the air shower density and its analysis using a CNN. In section 4, we applied the CNN method with the image data to the Crab Nebula observation and evaluated the γ/CR separation performance.

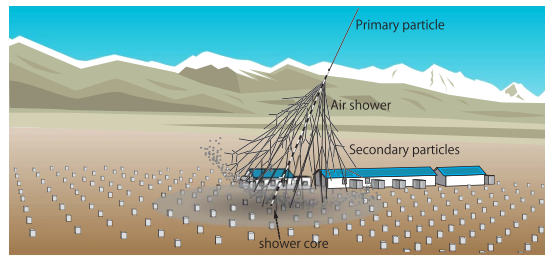## 2. Air shower observation with the Tibet air shower array

Primary higher-energy particles collide with atmospheric nuclei, producing an air shower that reaches the ground, as shown in figure 1(a). The main components of the particles reaching the ground in an air shower are electrons, gamma rays, muons, etc which are called secondary particles. The total number of charged secondary particles is approximately the fourth power of 10 for an incident particle of 10 TeV, and they spread over several hundred meters from the center of the shower [4]. One method of observing such air showers is to measure the number of secondary particles and their arrival times on a two-dimensional plane with a ground-based air shower array with radiation detectors installed several hundred meters square on the ground, which is the approach used in the TibetASγ experiment. Using the information from the detected secondary particles, the core position of the air shower ('shower core'), shower axis, and other parameters are reconstructed. 1(b) shows an example of an event display of an air shower event measured in the air shower array of the TibetASγ experiment. The size and color of each circle represent the logarithmic particle density and the relative hit timing in each detector, respectively. They defined the shower core using the center of gravity of the particle number density distribution in a two-dimensional plane and calculated the radial spread of secondary air shower particles ('lateral spread of secondary air shower particles') from the shower core. The energy of the air shower is determined from the sampled number of secondary particles falling into the detectors. From the hit timing of the secondary particles measured at each detector, the 'shower axis' is reconstructed in three dimensions and used as the primary particles' arrival direction.

The Tibet ASγ experiment, located at the Yangbajing Cosmic Ray Observatory in China (longitude $90°.522$ E, latitude $30°.102$ N, altitude 4300 m above sea level, atmospheric depth 606 g cm^{-2}), was undertaken to study gamma-ray astrophysics at energies greater than several TeV and to measure the abundance of nuclear species in cosmic rays ('chemical composition') with energies of several petaelectron volts.
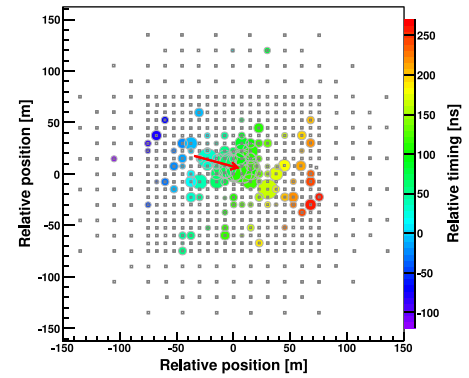
The first air shower array was constructed in 1990 [14]. After several upgrades and the construction of the MD, the current Tibet air shower array has a detection area of approximately 65 700 m^2 (figure 2). Each detector comprises a 3 cm thick plastic scintillator under a 5 mm thick lead plate with a detection area of 0.5 m^2, and photomultiplier tubes. The angular resolution is approximately $0.5°$ and $0.2°$ at 10 and 100 TeV gamma rays, respectively [1]. The energy resolution is approximately 40%, 20%, and 10% for 10, 100, and 400 TeV gamma rays, respectively [1, 15]. The systematic uncertainty in the absolute energy scale for ∼10 TeV is less than $\pm12\%$ [16]. Until 2014, observations were conducted only with the Tibet air shower array. In 2014, a water-Cherenkov-type MD was installed 2.4 m below the Tibet air shower array, enabling hybrid observations using two kinds of detectors. In the present study, we analyzed Monte Carlo data under the assumption that the Tibet air shower array consists of 597 plastic scintillator detectors (figure 2).

3. Separation methods for gamma-showers and hadronic-showers and their performance

We applied some separation method using particle number data per square meter ('particle number density') collected by each detector and then compared their performance. We first designed multiple feature values to represent the characteristics of lateral spread of secondary particles from the shower core and their azimuthal spread, which indicates the asymmetry of secondary particles around the shower core. We attempted the
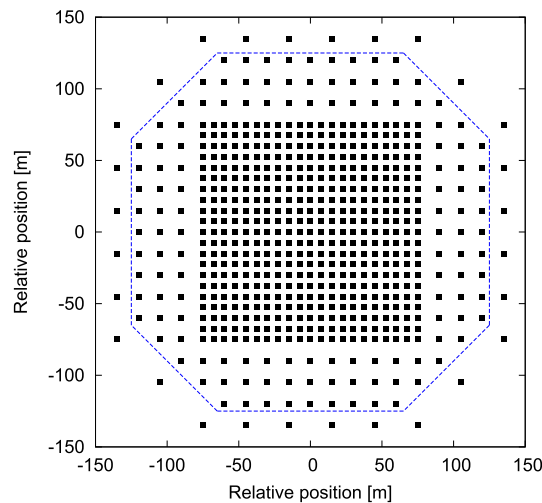
(a) Schematic diagram of air shower and observation experiment

(b) Event display of an air shower detected by the Tibet air shower array

Figure 1. (a) shows a schematic of air showers and observations of the Tibet ASγ experiment. Primary particles generate air showers through interactions with atmospheric nuclei. The secondary particles of the air shower, upon reaching the ground, are detected by an array of regularly spaced particle detectors. Using the information from the detected secondary particles, the shower core location, shower axis, and other parameters are reconstructed. (b) shows an event display of the gamma-showers of energy 114 TeV generated by a Monte Carlo simulation. Dots represent scintillation detectors. The size and color of each circle represent the logarithmic particle number density and the relative timing in each detector, respectively. The arrow head and direction indicate the shower core location and arrival direction, respectively.

Figure 2. Top view of the Tibet air shower array. The black squares represent surface $0.5\,\text{m}^2$ scintillation counters composing the Tibet air shower array. The air shower array, covering an area of $65\,700\,\text{m}^2$, consists of 597 scintillation counters. The area enclosed by the dashed blue line indicates the inner area used under selection condition (ii) of section 3.1.3.

separation using only each single feature value. We then applied an MLP-type neural network with multiple feature values and compared the separation performance with the results obtained using a single feature value. Finally, we applied a CNN method in which image data for the particle number density distribution collected by the Tibet air shower array are input to the neural network instead of feature values [11]. The MLP method has been applied in several previous air shower observation experiments, and its validity has been demonstrated for specific purposes. For instance, Tibet ASγ applied an MLP to the combined observation data obtained from the Tibet air shower array and the shower core detector for measuring particles in the shower core, leading to spectra of proton and helium as the main components of the elemental ratios of cosmic rays with energies in the petaelectron volt range [17–20]. This method requires the extraction of feature values from the particle density distribution in advance; however, it also requires clarification of which feature values are practical for separation. Therefore, finding the optimal set of feature values is critical.

In its computational process, the CNN automatically optimizes the parameters for extracting features from the image data. However, in the neural networks' learning process, the degrees of freedom of the network need to be appropriately matched to the degrees of freedom of the problem to yield correct results [21]. A CNN is highly capable of extracting features from data; however, its complex structure can

Table 1. Simulation settings.

Air shower simulation	CORSIKA (version 7.6400)
Primary energies	Gamma rays:10 TeV, 100 TeV
	Protons: 21 TeV, 165 TeV
Zenith angle	0°
Observation altitude	4300 m
Minimum kinetic energy	Electrons, photons, neutral π mesons: 1 MeV
of secondaries	Hadrons: 50 MeV, muons : 50 MeV
Interaction model	EPOS LHC v3400 and FLUKA 2011.2x
AS core position	$<$50 m from the array center
Detector simulation	GEANT4.10.02

lead to overtraining (overfitting) when the number of training data is insufficient. In such cases, the learned model becomes highly dependent on the training data, resulting in reduced classification performance for data other than the training data. In the current problem (γ/p-separation and γ/CR-separation), each dataset is an image of a particle number distribution detected by more than 500 detectors. Complex models are considered more optimal for extracting unknown valid features from them. However, because air shower events contain noise and fluctuations, using complex models can lead to overfitting due to insufficient training data. Thus, the degree of freedom of the CNN must be set appropriately.

To evaluate and compare the separation performance of each method, we carried out a Monte Carlo simulation for vertically incident gamma-showers and proton-showers with mono initial energies. Under these simplified simulation conditions, the degrees of freedom, such as the incident angle and the primary energy, are reduced and there is no influence of the chemical composition of cosmic rays. Consequently, separation becomes easier and each method can be more easily evaluated. In addition, because the classifier at the end of a CNN has the same type of structure as an MLP, we can evaluate the suitability of the network structure and the feature extraction capability of a CNN by comparing the separation performance of both methods.

3.1. Monte Carlo data of air showers measured with Tibet air shower array

First, we generated an air shower using the simulation code CORSIKA (version 7.6400) [22] and recorded information for secondaries, such as the particle ID, the position, and the arrival timing at Tibet altitude (4300 m). We then calculated the response of each scintillation detector to the charged particles and gamma rays using the Geant4.10.02 toolkit. Finally, we applied the same analysis methods as those used in the observation experiment to the particle densities and hit timings in each detector and reconstructed the energy and arrival direction of the primary particles. This procedure enabled the creation of simulated data in the same format as the experimental data.

3.1.1. Air shower generation

The Tibet air shower array was designed to measure gamma-showers and hadronic-showers in the energy range from a few TeV to approximately 10 PeV [23]. This study investigated the separation performance in the range from several tens of TeV to a few hundred TeV. To study each method's characteristics and compare their performance, we used vertically incident gamma-showers at 10 TeV and 100 TeV. Table 1 summarizes the configurations used in the Monte Carlo simulations. For the secondary particle energies, electrons, photons, and π^0 mesons were calculated down to 1 MeV, whereas other hadronic particles were calculated to 50 MeV. Muons were also calculated to 50 MeV. The recorded information included the particle energy, direction, arrival timing, and arrival position at an altitude of 4300 m.

We determined the primary energy using the sum of detected particle densities measured by each scintillator in the Tibet air shower array. The average values of the secondary particle densities for gamma-showers and hadronic-showers differ if the same primary energies are assumed because only part of the energy of primary hadrons are used to develop the electromagnetic showers through a neutral π meson production channel. Therefore, for 10 TeV gamma-showers, we generated proton-showers of 21 TeV to achieve a similar number of detected particles and separated those showers (10 TeV-γ/p-separation). For 100 TeV gamma-showers, proton-showers of 165 TeV were generated (100 TeV-γ/p-separation). This study used the EPOS LHC model [24] for calculations in the high-energy range and the FLUKA model [25] in the low-energy regime; these models have been widely used in previous studies.

3.1.2. Detector response

Each detector had a 3 cm thick plastic scintillator with an area of 0.5 m^2 and a 5 mm thick lead plate placed on the top; a total of 597 detectors were arranged mainly at 7.5 m intervals (figure 2) [26]. To avoid a

systematic bias caused by different spacings between detectors used in the Tibet air shower array, the position of the simulated shower cores was limited within a radius of 50 m from the array center, where 7.5 m spacing is used exclusively. Using the GEANT4.10.02 code, we calculated the energy loss and the hit timing responses of secondary particles within the plastic scintillator for each of the charged particles in an air shower generated with CORSIKA.

3.1.3. Air shower reconstruction

The measured energy loss in each scintillation detector was converted into the corresponding particle number, and the sum of particles for all detectors was approximately proportional to the energy of primary gamma rays and primary cosmic rays. We used the number of particles per square meter (particle number density; ρ [particle m^{-2}]) and the sum of ρ ($\Sigma\rho$) so that we could treat all detectors the same, even when detectors with different areas were mixed. The shower reconstruction used the hit timing to estimate the arrival direction [1]. In the present study, we imposed the following selection conditions (i)–(v) to use well-reconstructed air shower events from the detected data.

(i) At least four detectors must be hit within the time width of coincidence of 600 ns.
     This condition is the trigger condition in data acquisition. Secondary air shower particles are nearly as fast as light; most reach the ground within tens of nanoseconds. This condition is necessary for air showers to be measured selectively in actual experiments.
(ii) The number of detectors with $\rho \geqslant 0.6\,\mathrm{m}^{-2}$ must be at least four in the inner detectors of the array (refer to the caption of figure 2).
(iii) Five or more of the top six detectors with the most detected particles must be contained in the inner detectors.
(iv) The location of the shower core determined by the analysis must be within a 50 m radius from the array center.
     We determined the direction of arrival of the air shower by fitting a reverse-conic-type shape to the front surface of the shower particles, which was calculated using the hit timing of each detector [26–28].
(v) The residual error, which represents the goodness of fit determined by this method, should be less than 1.0 m.

After analyzing the MC data for the above selection conditions, we obtained the $\Sigma\rho$ distributions of 10 TeV (21 TeV) and 100 TeV (165 TeV) gamma-showers (proton-showers), as shown by the red (green) histograms in figure 3. To prevent differences in the shape of the $\Sigma\rho$ distribution between gamma-showers and proton-showers from influencing the separation, for the γ/p-separation analysis, we reduced the proton events shown in the green histograms in figure 3 to have the same number of events as gamma events at each $\Sigma\rho$ bin. Specifically, the same number of gamma-showers and proton-showers whose $\Sigma\rho$ distribution matches each other were used in the following analysis.

3.2. Separation with single feature value

We tested the following six feature values representing the characteristics of the spread of secondary particles in addition to $\Sigma\rho$ and compared the separation performance achieved with each feature value.
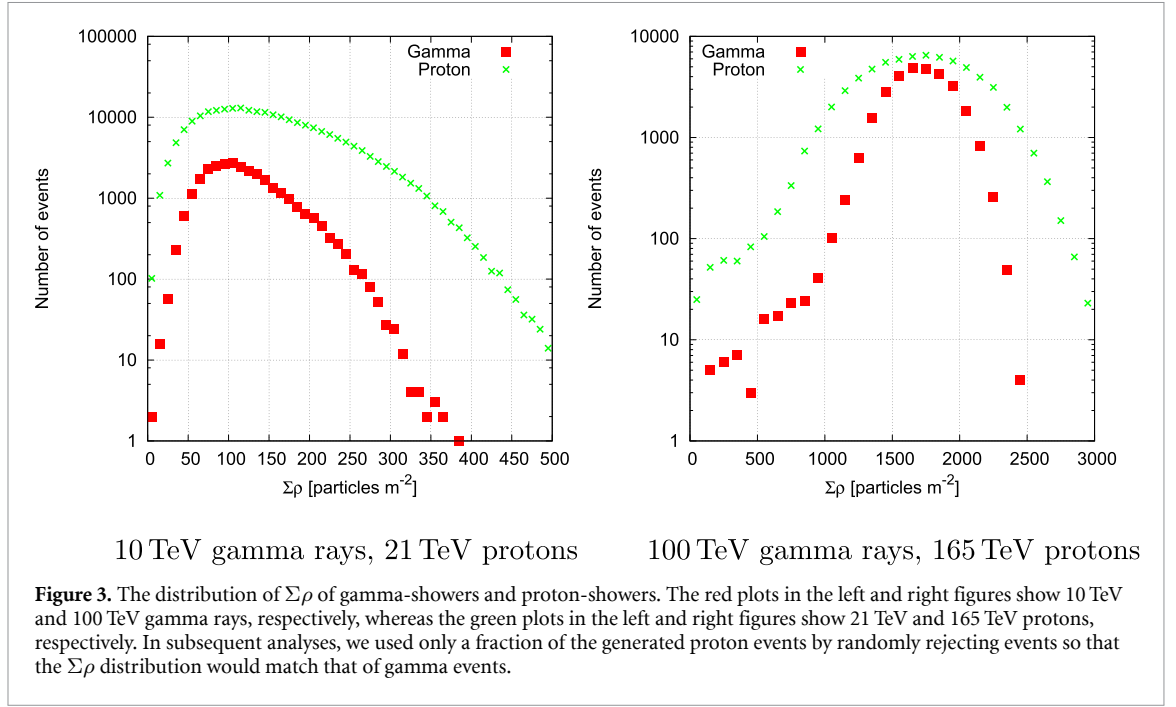
3.2.1. Feature value of lateral spread

The first choice as the feature value of air showers is the average lateral spread of the secondary particles. To calculate this value, we used the following equation, which is the same type of equation as the Tibet ASγ used in a previous study of the chemical composition of cosmic rays [29, 30].

$$R = \frac{\sum_i \rho_i \times r_i}{\sum_i \rho_i},\qquad(1)$$

where ρ_i is the particle number density detected by the ith detector and r_i is the distance of the ith detector from the shower core position.

Figure 4(a) shows the distributions of R for 10 TeV gamma-shower events (red points) and 21 TeV proton-shower events (green points), respectively. Figure 5(a) shows the distributions for 100 TeV gamma-shower events and 165 TeV proton-shower events. Gamma-showers exhibit a higher concentration of secondary particles near the shower core, resulting in a smaller value of R. However, proton-showers consist of multiple electromagnetic showers developed from the decay of π^0 mesons produced by nuclear interactions in the atmosphere and having an average transverse momentum of several hundred MeV/c, resulting in a larger value of R.

10 TeV gamma rays, 21 TeV protons   100 TeV gamma rays, 165 TeV protons

Figure 3. The distribution of $\Sigma\rho$ of gamma-showers and proton-showers. The red plots in the left and right figures show 10 TeV and 100 TeV gamma rays, respectively, whereas the green plots in the left and right figures show 21 TeV and 165 TeV protons, respectively. In subsequent analyses, we used only a fraction of the generated proton events by randomly rejecting events so that the $\Sigma\rho$ distribution would match that of gamma events.

3.2.2. Feature value of concentration

We attempted to extract the multicore structure in proton-showers as the next feature value. We define values

$$N_j = \rho_j \times r_j, \tag{2}$$

where ρ_j represents the jth largest value of particle density within a shower whose distance from the shower core is denoted by r_j. In the case of 100 TeV gamma-showers and 165 TeV proton-showers, we used only detectors located at distances greater than 50 m from the shower core because the particle density near the shower core is high for both showers and the difference between them is not substantial.

Figures 4(b)–(d) show the distributions of N_j ($j = 1, 2, 3$) for 10 TeV gamma-showers and 21 TeV proton-showers, and figures 5(b)–(d) show the distributions for 100 TeV gamma-showers and 165 TeV proton-showers. The N_j value for the gamma-showers tends to be smaller than that for the proton-showers because of the higher concentration of secondary particles near the core.
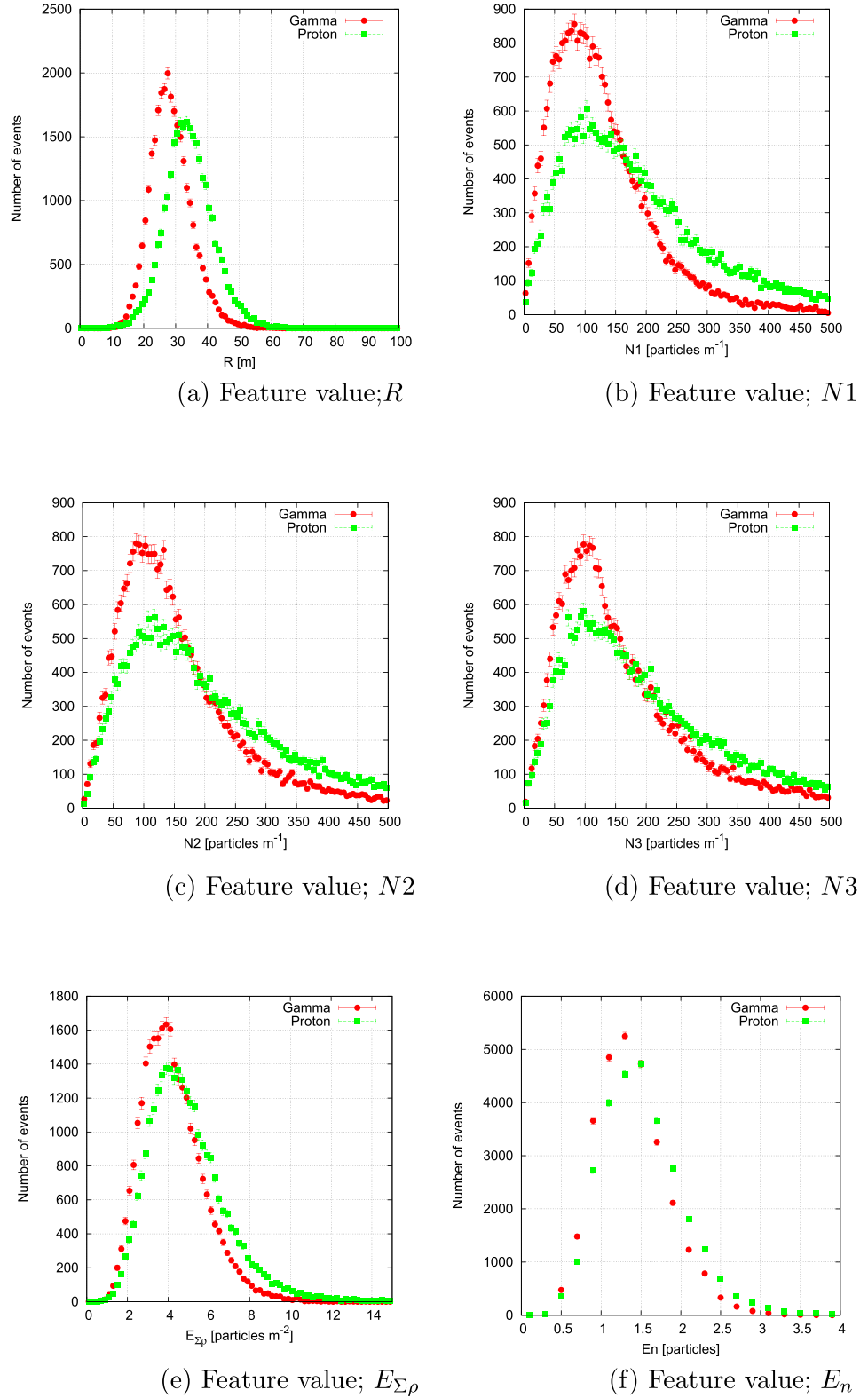
3.2.3. Feature value of asymmetry

The fluctuations in the generation altitude of π^0 mesons can affect the density distribution of secondary particles at the observation site, resulting in a circumferential asymmetric distribution. We therefore divided the circular region around the shower core into eight fan-shaped segments (figure 6). We defined the feature values using the particle number density and the number of hit detectors in each segment. We then defined $E_{\Sigma\rho}$ and E_n as feature values by the following equations:

$$E_{\Sigma\rho} = \sqrt{\frac{1}{8}\sum_{k=1}^{8}\left\{\left(\sum\rho\right)_k - \frac{\left(\sum\rho\right)_{all}}{8}\right\}^2}, \tag{3}$$

$$E_n = \sqrt{\frac{1}{8}\sum_{k=1}^{8}\left(n_k - \frac{n_{all}}{8}\right)^2}, \tag{4}$$

where $\left(\sum\rho\right)_k$ represents the sum of particle number density in the kth region out of the eight regions, $\left(\sum\rho\right)_{all}$ represents the sum of the particle number density in all regions, n_k represents the number of hit detectors in the kth region, and n_{all} represents the total number of hit detectors in all regions. To avoid the influence of core position fluctuations, especially near the shower core, we used only detectors located at distances greater than 30 m and 50 m from the shower core for the 10 TeV- and 100 TeV-γ/p-separation, respectively. The average numbers of hit detectors selected by these criteria are 18.5

Figure 4. Distribution of feature values used to separate 10 TeV gamma-showers and 21 TeV proton-showers. The red plots show gamma-ray events, and the green plots show proton events.

for 10 TeV gamma-showers, 21.2 for 21 TeV proton showers, 95.4 for 100 TeV gamma-showers, and 106 for 165 TeV proton-showers.

Figures 4(e) and (f) show the distributions of $E_{\Sigma\rho}$ and E_n, respectively, for 10 TeV-γ/p-separation. Figures 5(e) and (f) show the same for 100 TeV-γ/p-separation. As expected, the proton-showers tend to have slightly larger values of $E_{\Sigma\rho}$ and E_n compared with the gamma-showers.
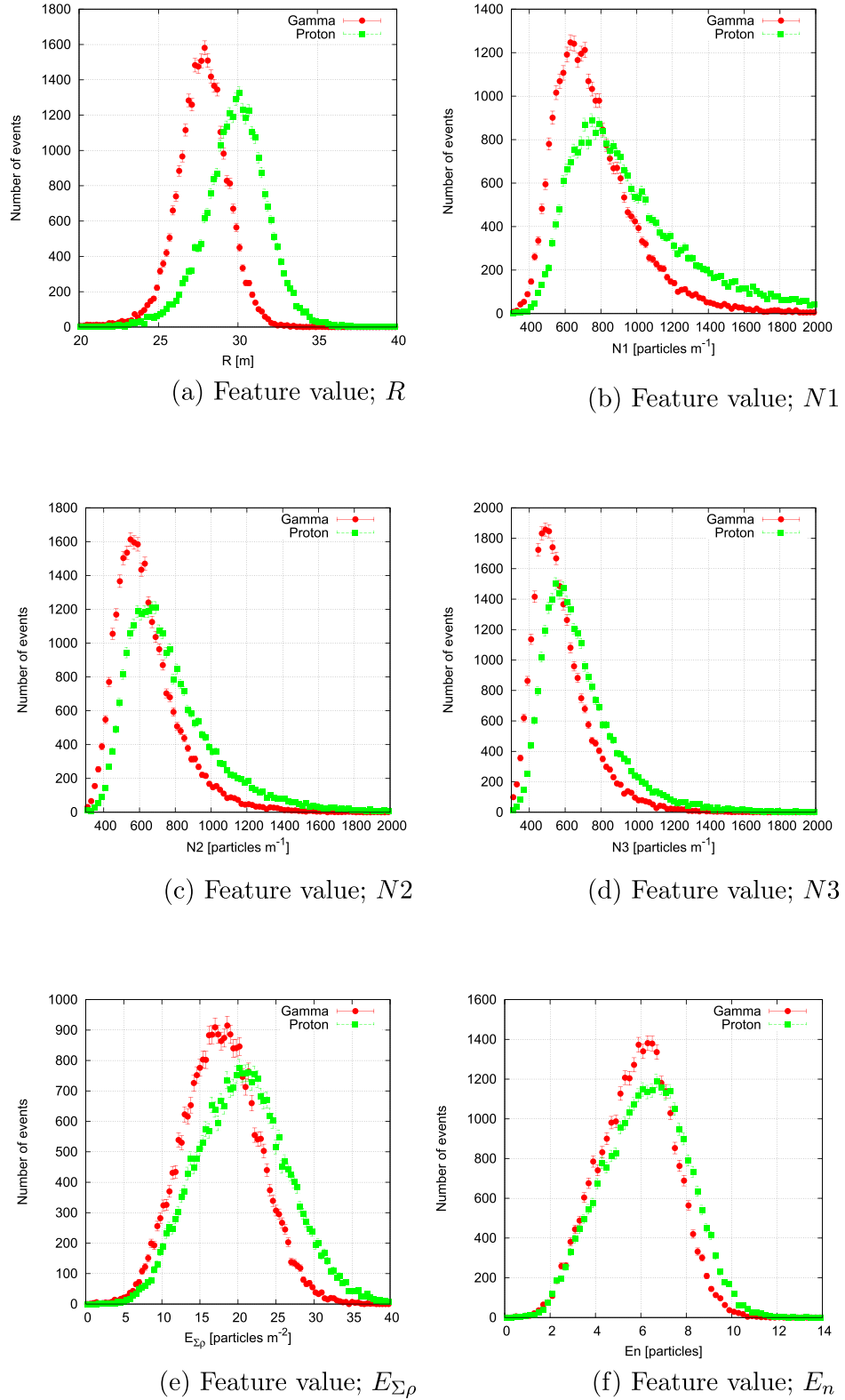
Figure 5. Distribution of feature value used to separate 100 TeV gamma-showers and 165 TeV proton-showers. The red plots show gamma-ray events, and the green plots show proton events.

3.2.4. Separation performance with a single feature value

To evaluate the separation performance using a single feature value described in section 3.2, we used the receiver operating characteristic (ROC) curves and the area under the curve (AUC) values [31]. Figure 7 shows the ROC curves for each feature value, where the *x*-axis represents the false positive rate (i.e. the ratio of proton events incorrectly classified as gamma-ray events) and the *y*-axis represents the true positive rate (i.e. the ratio of gamma-ray events correctly classified as gamma-ray events). The error bars in each plot

Figure 6. An example of eight fan-shaped segments for calculating $E_{\Sigma\rho}$ and E_n values. This figure is the result of a vertically incident 100 TeV gamma-shower. Open squares represent scintillation detectors. The color of each rectangle represents the particle number density, and each color is mapped by $\log_{10}(\rho + 0.01)$. Each segment is divided by black lines with the intersection of the center of the shower core. The black circle represents a distance of 50 m from the shower core location, and the hit detectors outside this circle were used in the calculation.
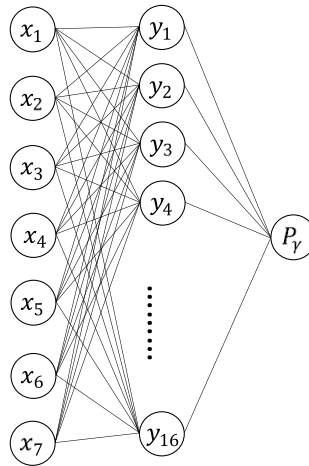

(a) 10 TeV gamma rays, 21 TeV protons   (b) 100 TeV gamma rays, 165 TeV protons

Figure 7. ROC curves for separation with a single feature value. (a) shows the results for 10 TeV-γ/p-separation, and (b) shows the results for 100 TeV-γ/p-separation. The black plots show the result of separation with R, the red plot with N_1, the blue plot with N_2, the green plot with N_3, the cyan plot with by $E_{\Sigma\rho}$, the orange plot with E_n, and the brown plot with $\Sigma\rho$.

represent the standard deviation of the binomial distribution for both the vertical and horizontal axes. The left figure represents the results of 10 TeV-γ/p-separation, and the right figure represents the results of 100 TeV-γ/p-separation. The black, red, blue, green, cyan, orange, and brown plots correspond to the separations based on R, N_1, N_2, N_3, $E_{\Sigma\rho}$, E_n, and $\Sigma\rho$, respectively. The separation based on R is more effective than the separations based on the other values. Table 2 presents the AUC values calculated from the ROC curves for each feature value. We defined the errors in the AUC values as the upper (lower) error, which changes the AUC value when all the plot points on the ROC curve shift in the upper-left (lower-right) direction according to the error bars. At both energies, the AUC value was highest when the separation was performed using R with a value of 0.701 for 10 TeV-γ/p-separation and 0.808 for 100 TeV-γ/p-separation. In addition, as the energy increases, the AUC values tend to increase except for E_n and as expected $\Sigma\rho$, indicating their poorer characteristics compared with the other feature values. The average numbers of hit detectors are 35.3 and 36.2 for 10 TeV gamma-showers and 21 TeV proton-showers, respectively, and 206.4 and 213.2 for 100 TeV gamma-showers and 165 TeV proton-showers, respectively. These results show that the increase in the number of hit detectors along with increasing energy led to a clearer characterization of the air shower's secondary particles, improving separation performance.

Table 2. AUC with single feature value.

Feature value	10 TeV-γ/p-separation	100 TeV-γ/p-separation
R	$0.701^{+0.011}_{-0.011}$	$0.808^{+0.009}_{-0.009}$
N_1	$0.653^{+0.011}_{-0.011}$	$0.691^{+0.011}_{-0.011}$
N_2	$0.624^{+0.011}_{-0.011}$	$0.678^{+0.011}_{-0.011}$
N_3	$0.588^{+0.011}_{-0.011}$	$0.668^{+0.011}_{-0.011}$
$E_{\Sigma\rho}$	$0.626^{+0.011}_{-0.011}$	$0.645^{+0.011}_{-0.011}$
E_n	$0.596^{+0.011}_{-0.011}$	$0.570^{+0.012}_{-0.011}$
$\Sigma\rho$	$0.539^{+0.012}_{-0.012}$	$0.529^{+0.012}_{-0.012}$

Figure 8. The architecture of MLP, which has seven inputs (x_i ($i = 1, 2, \ldots, 7$)), 16 nodes (y_i ($i = 1, 2, \ldots, 16$)) in a middle layer, and one output (P_γ).

3.3. Separation with an MLP with multifeature values

We investigated the γ/p-separation performance of an MLP neural network. We used the seven feature values defined in section 3.2 (i.e. $R, N_1, N_2, N_3, E_{\Sigma\rho}$, and E_n) and the sum of the particle number density detected by the Tibet air shower array ($\Sigma\rho$) as the inputs for an MLP model.
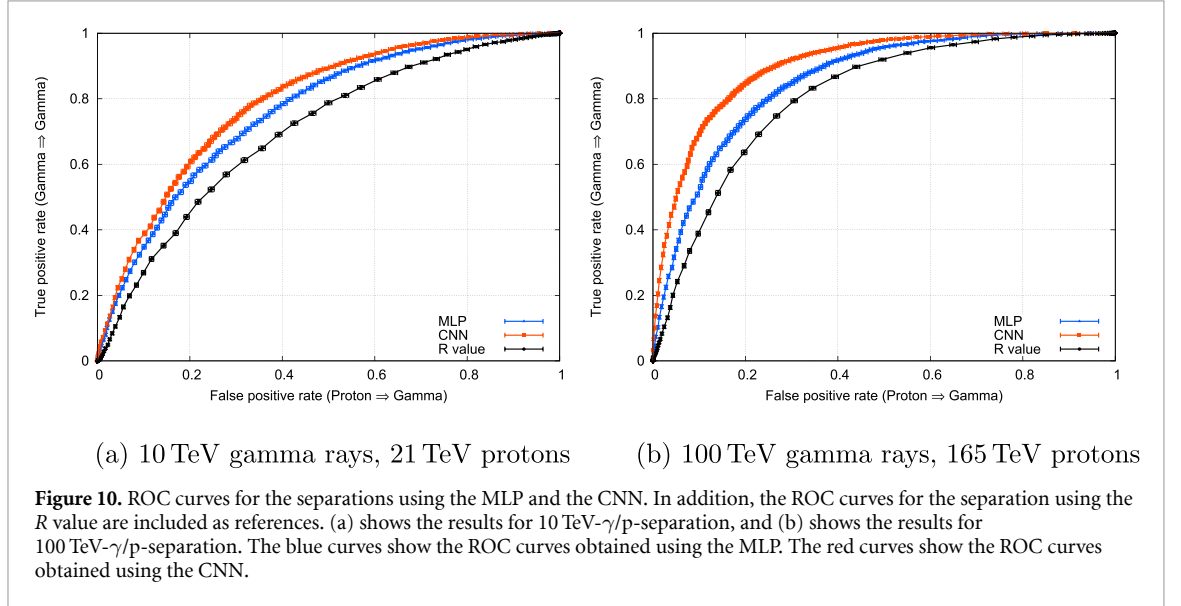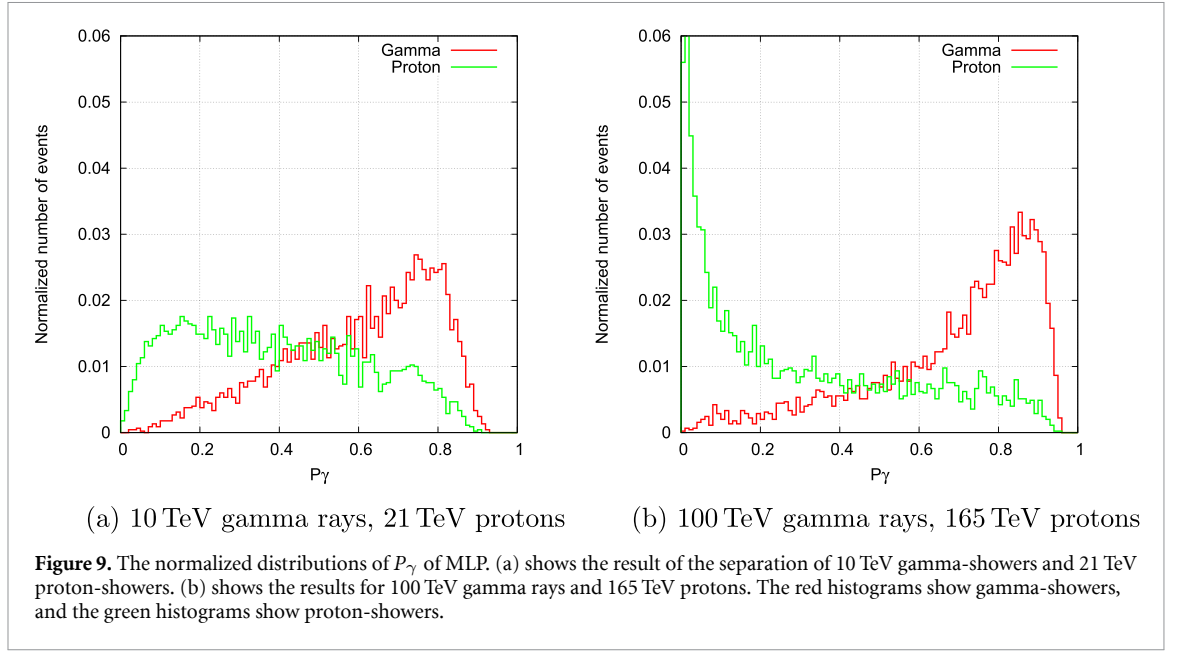
3.3.1. MLP model and analysis procedure

This study used the Keras neural network library [32] to test an MLP model with seven input dimensions, one hidden layer with 16 nodes, and one output dimension (figure 8). To input the seven feature values into the MLP, we standardized each feature value as follows:

$$\frac{x_{i,j} - \mu_i}{\sigma_i} \rightarrow x_{i,j}, \tag{5}$$

where μ_i ($i = 1, 2, 3, 4, 5, 6, 7$) represents the mean of the ith feature value, given by $\mu_i = \frac{1}{N}\sum_{j=1}^{N} x_{i,j}$; σ_i^2 represents the variance, defined as $\sigma_i^2 = \frac{1}{N}\sum_{j=1}^{N}(x_{i,j} - \mu_i)^2$; N denotes the total number of gamma-showers and proton-showers in the training dataset; and $x_{i,j}$ ($i = 1, 2, \ldots, 7; j = 1, 2, \ldots, j_{\max}$) represents the standardized feature value, where j refers to the jth event out of N events and i refers to the ith feature value. For the training data, we prepared 24 000 events each for gamma-showers and proton-showers, resulting in $N = 48\,000$ and $j_{\max} = 48\,000$. Through this standardization process, we set the mean of each feature value to 0 and the variances to 1 with equation (5).

These inputs were connected to a densely connected hidden layer of 16 nodes with sigmoid-type activation functions. The sigmoid function was also used as the output function, and the output value (P_γ) was calculated. For training validation, we used 10% of the events from the entire training dataset and computed the validation loss using cross entropy at each epoch of the training process. We obtained a trained classifier by minimizing the validation loss. To evaluate the γ/p-separation performance, we calculated P_γ for 4500 gamma-showers and proton-showers for each.
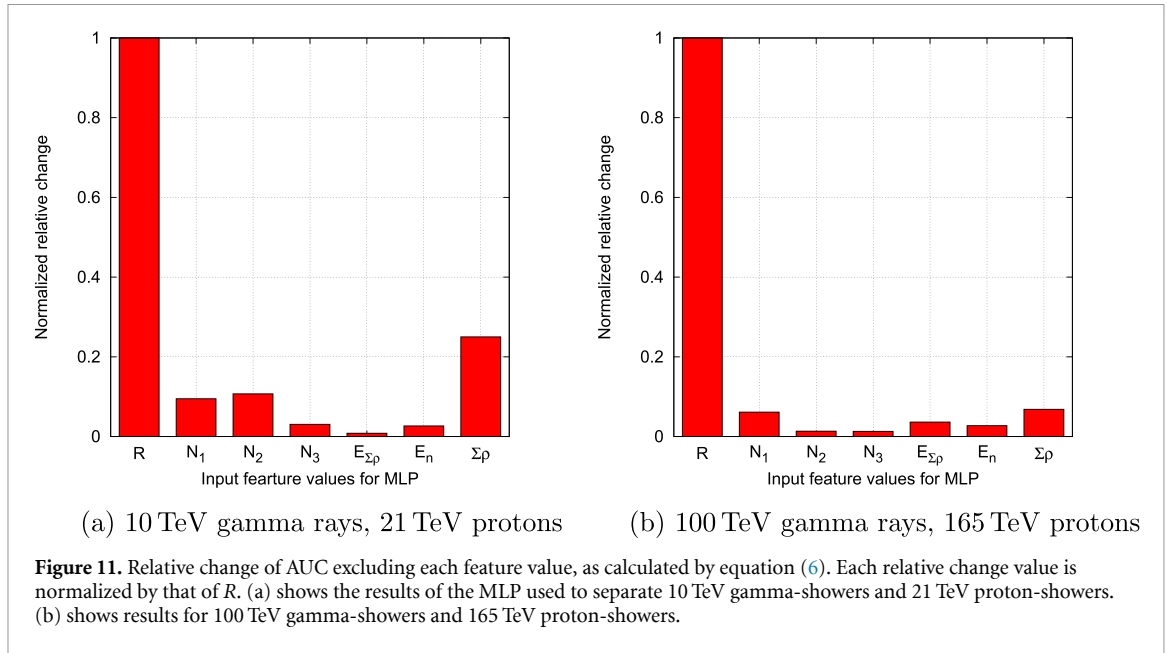
(a) 10 TeV gamma rays, 21 TeV protons   (b) 100 TeV gamma rays, 165 TeV protons

Figure 9. The normalized distributions of P_γ of MLP. (a) shows the result of the separation of 10 TeV gamma-showers and 21 TeV proton-showers. (b) shows the results for 100 TeV gamma rays and 165 TeV protons. The red histograms show gamma-showers, and the green histograms show proton-showers.

(a) 10 TeV gamma rays, 21 TeV protons   (b) 100 TeV gamma rays, 165 TeV protons

Figure 10. ROC curves for the separations using the MLP and the CNN. In addition, the ROC curves for the separation using the R value are included as references. (a) shows the results for 10 TeV-γ/p-separation, and (b) shows the results for 100 TeV-γ/p-separation. The blue curves show the ROC curves obtained using the MLP. The red curves show the ROC curves obtained using the CNN.

3.3.2. Separation performance with an MLP

Figures 9(a) and (b) show P_γ distributions. The red histograms in figures 9(a) and (b) show P_γ distributions of the gamma-showers, and the green histograms represent the proton-showers. The P_γ will be close to 1 if the separated events are gamma-ray-like and close to 0 if they are proton-like. The P_γ distribution of 10 TeV gamma-showers reaches the maximum at ∼0.8 (figure 9(a)). If we set a threshold of $P_\gamma = 0.5$ for separation, we can retain 71% of gamma-showers while excluding 63% of proton showers. For 100 TeV gamma-showers (figure 9(b)), the distribution of P_γ reaches the maximum at ∼0.9 and the proton showers are concentrated around 0. With a threshold of $P_\gamma = 0.5$, the remaining rate of gamma-showers is 81% and the rejection rate of proton showers is 73%.

To evaluate the γ/p-separation performance, we calculated the ROC curves and AUC values from the P_γ distributions (figures 10(a) and (b)). The blue plots represent the results for 10 TeV-γ/p-separation and 100 TeV-γ/p-separation. The AUC values are AUC $= 0.761^{+0.010}_{-0.019}$ for 10 TeV-γ/p-separation and AUC $= 0.854^{+0.008}_{-0.008}$ for 100 TeV-γ/p-separation. Thus, the separation performance improves as the energy increases. The MLP results in an improvement of ∼5% in the AUC value compared the results obtained using only R (section 3.2.4).

To understand the rationale behind the MLP classification, we compared the AUC values in the following procedure to examine the significance of seven feature values. First, we changed only the ith feature value (x_i)

(a) 10 TeV gamma rays, 21 TeV protons   (b) 100 TeV gamma rays, 165 TeV protons

Figure 11. Relative change of AUC excluding each feature value, as calculated by equation (6). Each relative change value is normalized by that of *R*. (a) shows the results of the MLP used to separate 10 TeV gamma-showers and 21 TeV proton-showers. (b) shows results for 100 TeV gamma-showers and 165 TeV proton-showers.

of the seven feature values to zero and calculated the AUC (AUC_i) using the modified feature values and the trained MLP. We then calculated the relative change $\mathcal{R}_i$ as

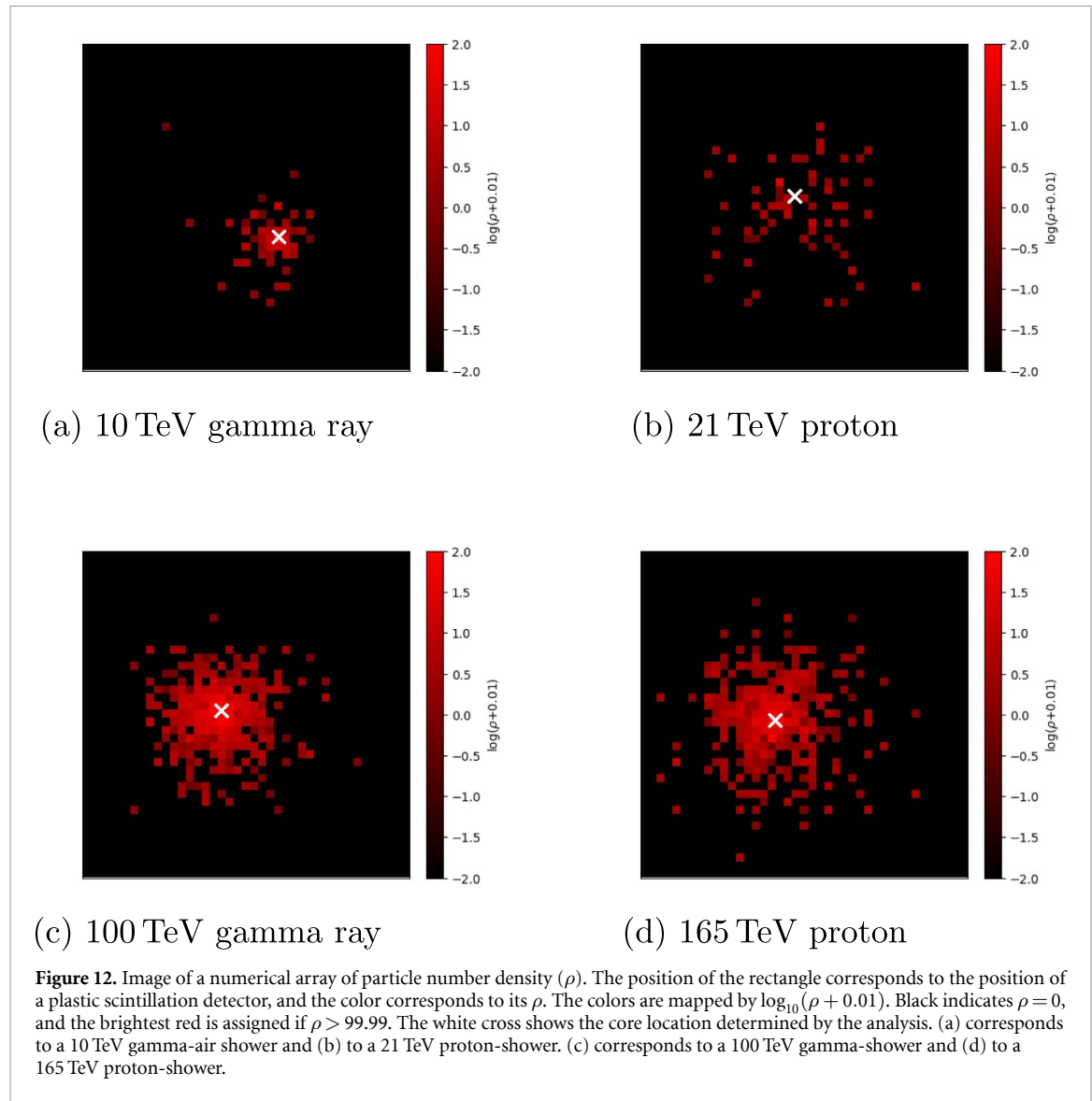$$\mathcal{R}_i = \frac{AUC_{all} - AUC_i}{AUC_{all}}, \tag{6}$$

where AUC_{all} represents the AUC value obtained using all the feature values. Figures 11(a) and (b) show the $\mathcal{R}_i$ for each feature value. In these graphs, each value of $\mathcal{R}_i$ is normalized with that of *R*. Figure 11(a) represents the results for 10 TeV γ/p-separation. The relative change for *R* is more than four times larger than the relative changes for the other feature values. In addition, the distribution of $\Sigma\rho$ was matched between gamma-showers and proton-showers. Therefore, $\Sigma\rho$ cannot be used alone for γ/p-separation. However, its relative change values are second in magnitude to that of *R*. The aforementioned results show that the MLP separates events by combining $\Sigma\rho$ with other feature values in the computational process. Figure 11(b) represents the results for 100 TeV γ/p-separation. The relative change of *R* is more than ten times larger than the relative changes of the other feature values, and the significance of the other feature values were equally smaller. These results indicate that the lateral spread of secondary particles is critical for γ/p-separation and that the practical feature values depend on energy. Also, all $\mathcal{R}_i$ in figures 11(a) and (b) are positive, indicating that none of the feature values are useless.

3.4. Separation with a CNN with image-like data of particle density distribution
In our previous work [11], we also attempted γ/p-separation using a CNN commonly used for image recognition. Because it automatically optimizes filter parameters for feature extraction from images. We therefore expected improved separation performance.

3.4.1. Design of image-like data
We created 2D image-like data from the positions of each detector and the particle density [11]. To convert the particle-density data from 597 detectors arranged in a grid pattern into numerical array data in image format, we divided the range of 307.5×307.5 m into a 41×41 grid and stored the value of detected particle density in each element of the numerical array. We set the value to zero for array elements corresponding to positions without detectors. In addition, to use a CNN designed for RGB color image input, where each pixel consists of three components, we stored only one component of the particle density and zeros for the other two components. This procedure created numerical array data in image format with dimensions of $41 \times 41 \times 3$, representing the spread of the particle density. Figure 12 presents examples of numerical array data visualized as images. Figures 12(a)–(d) show the results for a 10 TeV gamma-shower, a 21 TeV proton-shower, a 100 TeV gamma-shower, and a 165 TeV proton-shower, respectively. The color of each rectangle represents the particle density of the corresponding detector. In addition, the white cross indicates the shower core position determined by the conventional analysis. In these examples, gamma-showers exhibit high density near the core and a symmetrical tendency for the particle density. By comparison, proton showers show high density near the core but an uneven surrounding density.

(a) 10 TeV gamma ray

(b) 21 TeV proton

(c) 100 TeV gamma ray

(d) 165 TeV proton

Figure 12. Image of a numerical array of particle number density (ρ). The position of the rectangle corresponds to the position of a plastic scintillation detector, and the color corresponds to its ρ. The colors are mapped by $\log_{10}(\rho + 0.01)$. Black indicates $\rho = 0$, and the brightest red is assigned if $\rho > 99.99$. The white cross shows the core location determined by the analysis. (a) corresponds to a 10 TeV gamma-air shower and (b) to a 21 TeV proton-shower. (c) corresponds to a 100 TeV gamma-shower and (d) to a 165 TeV proton-shower.

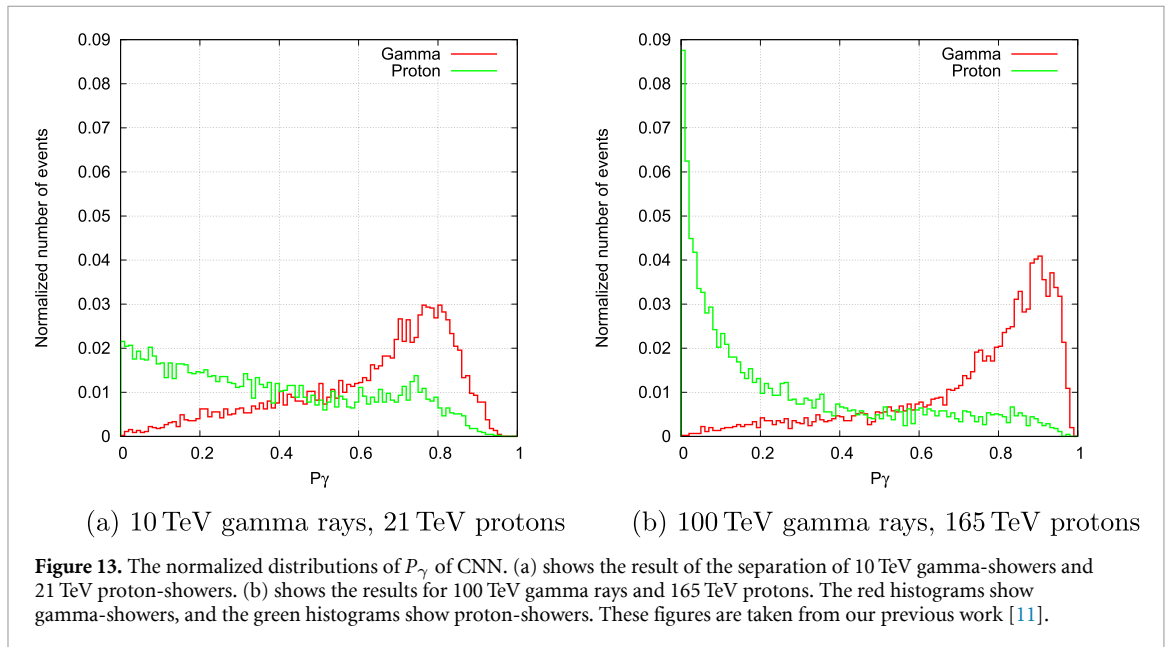*3.4.2. CNN model and analysis procedure*

We constructed a CNN architecture consisting of consecutive processes, referencing the Inception-v3 image recognition model [11, 33] and implementing it using the Keras interface [32]. The architecture is shown in table 3, where the first column represents the processing method and the second column indicates the input data size or the data size after each process. To match the input data size for the Inception-v3 model $(299 \times 299 \times 3)$, we enlarged the image-like data size to $(246 \times 246 \times 3)$—a sixfold increase. The increased components were filled with the same values as the original components. In the first step, features of image-like data were extracted using a convolutional layer with a filter size of (3×3) and a stride of one. In addition, in this layer, to prevent the data size from becoming too small, zero-padding was used to maintain the output size to be the same as the input size. We applied a dropout process with a dropout rate 0.1 to prevent overfitting [34]. This consecutive process was performed three times. Next, we reduced the data size using a max pooling layer with a size of (2×2) and a stride of two. We then performed additional image convolutions and extracted the features using three types of Inception modules shown in figures 5–7 of [33]. We subsequently calculated the average value for each channel of a (30×30) tensor and converted the tensor-format image data into a vector format using a global average pooling layer [35]. We performed classification based on the extracted features using a densely connected layer and applied a dropout process with a dropout rate of 0.4. Finally, the data were passed through a sigmoid output function to calculate the gamma-likeness P_γ value.

We analyzed the same air shower events used in the MLP method, using 10% of the training data as validation and calculating the validation loss at each epoch. The training results from the epoch with the lowest validation loss were used for the following test.
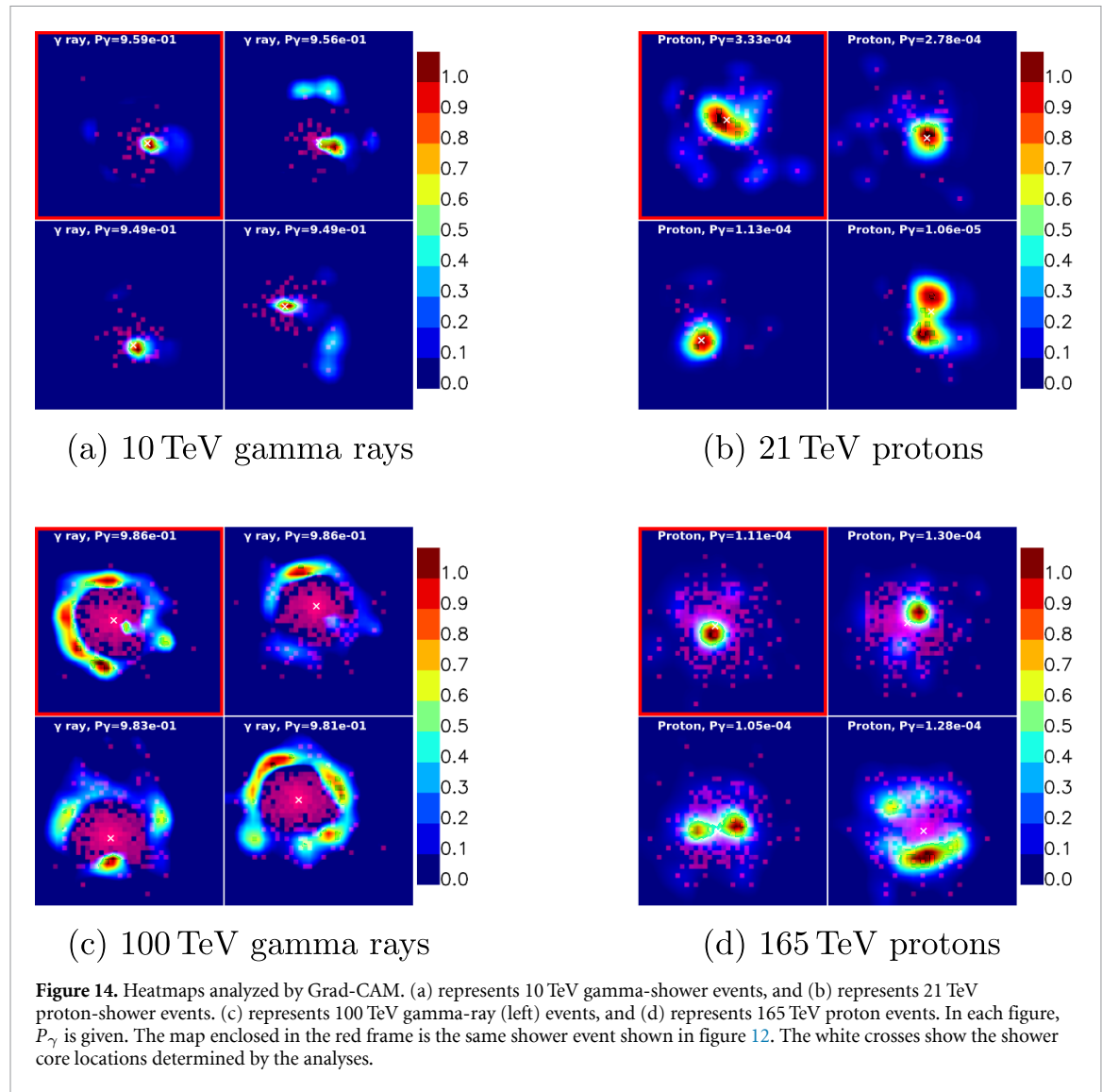
Table 3. Outline of the CNN architecture.

Procedure		Data size
Inputs	—	$246 \times 246 \times 3$
Convolution	filter size: 3×3, stride: 1	$246 \times 246 \times 32$
Dropout	dropout rate: 0.1	$246 \times 246 \times 32$
Convolution	filter size: 3×3, stride: 1	$246 \times 246 \times 32$
Dropout	dropout rate: 0.1	$246 \times 246 \times 32$
Convolution	filter size: 3×3, stride: 1	$246 \times 246 \times 64$
Dropout	dropout rate: 0.1	$246 \times 246 \times 64$
Max pooling	size: 2×2, stride: 2	$123 \times 123 \times 64$
Inception $\times 3$	see figure 5 in [33]	$123 \times 123 \times 128$
Max pooling	size: 2×2, stride: 2	$61 \times 61 \times 128$
Inception $\times 5$	see figure 6 in [33]	$61 \times 61 \times 256$
Max pooling	size: 2×2, stride: 2	$30 \times 30 \times 256$
Inception $\times 2$	see figure 7 in [33]	$30 \times 30 \times 510$
Global average pooling	—	510
Dense	—	200
Dropout	dropout rate: 0.4	200
Sigmoid classifier	—	1

(a) 10 TeV gamma rays, 21 TeV protons   (b) 100 TeV gamma rays, 165 TeV protons

Figure 13. The normalized distributions of P_γ of CNN. (a) shows the result of the separation of 10 TeV gamma-showers and 21 TeV proton-showers. (b) shows the results for 100 TeV gamma rays and 165 TeV protons. The red histograms show gamma-showers, and the green histograms show proton-showers. These figures are taken from our previous work [11].

3.4.3. Separation performance with the CNN

The red and green histograms in figure 13(a) show the P_γ distributions of 10 TeV gamma-showers and 21 TeV proton-showers. Figure 13(b) shows the results for 100 TeV gamma-showers and 165 TeV proton-showers. Most P_γ of 10 TeV gamma-showers have a value greater than 0.5, and the peak is at $\sim$0.8. By contrast, the maximum value for the P_γ distribution for 21 TeV proton-showers is in the region $P_\gamma < 0.5$. If we separate events with a threshold of $P_\gamma = 0.5$, we can reduce the proton-showers to 36% with 78% of gamma-showers remaining. The distribution of 100 TeV gamma-shower events has a maximum at $\sim$0.9, and most of the P_γs of protons are approximately zero. The separation with $P_\gamma = 0.5$ can reduce the proton-showers to 20% with 85% of gamma-showers remaining.

We also show the ROC curves of the CNN in figures 10(a) and (b). The red plots represent the results for 10 TeV-γ/p-separation and 100 TeV-γ/p-separation. The ROC values of the CNN were larger than those of the MLP. The AUC value for 10 TeV-γ/p-separation is $0.781^{+0.009}_{-0.009}$ and that for 100 TeV-γ/p-separation is $0.901^{+0.006}_{-0.006}$. The AUC values of the CNN are approximately five percent higher than those of the MLP, indicating superior separation performance. The CNN automatically extracts features, whereas the MLP utilizes feature values manually defined by us. Nonetheless, the CNN exhibits better separation performance. Consequently, we expect that, compared with an artificial method, the CNN has superior ability to extract suitable feature values.

Figure 14. Heatmaps analyzed by Grad-CAM. (a) represents 10 TeV gamma-shower events, and (b) represents 21 TeV proton-shower events. (c) represents 100 TeV gamma-ray (left) events, and (d) represents 165 TeV proton events. In each figure, P_γ is given. The map enclosed in the red frame is the same shower event shown in figure 12. The white crosses show the shower core locations determined by the analyses.

Knowing which physical quantities effectively contribute to separation in the analysis of the CNN is difficult. Still, multiple methods are available to identify which parts of the image-like data contribute to the separation. We used a technique known as Grad-CAM [36] to generate heatmaps that show the magnitude of the contribution of each element in the feature map (array data immediately before the global average pooling layer) to the final output of the CNN. In addition, because the feature map retains the positional information of the input image-like data, examining these heatmaps reveals which parts of the input data's features strongly affect the separation results. Figures 14(a)–(d) show the heatmaps for typical examples whose particle types are correctly identified. In addition, the maps enclosed in red frames correspond to the heatmaps of the shower events shown in figure 12. The reddish regions in the heatmaps indicate areas that contributed significantly to the separation. Figure 14(a) represents the results of a 10 TeV gamma-shower, with a P_γ value ranging from 9.49×10^{-1} to 9.59×10^{-1}. The core region appears prominently in red, indicating its substantial contribution to separation in these maps. Figure 14(b) represents the results of a 21 TeV proton-shower, with a P_γ value ranging from 1.06×10^{-5} to 3.33×10^{-4}. In addition to the core region, a wide range of surrounding areas is also a lighter shade of red, suggesting that multiple regions contributed significantly to the separation. Figure 14(c) represents a 100 TeV gamma-shower with a P_γ value ranging from 9.81×10^{-1} to 9.86×10^{-1}. In this case, the peripheral regions contribute more significantly to the separation than the core region. Figure 14(d) represents a 165 TeV proton-shower with a P_γ value ranging from 1.05×10^{-4} to 1.30×10^{-4}. These maps indicate both the core region and peripheral localized regions are utilized in the separation. These heatmaps thus indicate that feature extraction is energy-dependent. Gamma-showers have a higher particle density near the core than proton-showers with the same $\Sigma\rho$. At higher energies, however, the surrounding region becomes more important because, in both cases, many of the detectors near the core exhibit very high particle densities that are difficult to distinguish.

Table 4. CORSIKA-simulation settings for training and test data.

Primaries	Gamma rays	Cosmic rays
Energy spectra of primaries	$\frac{dN}{dE} \propto E^{-2.6}$	Shibata model [37]
Energy regions	300 GeV–10 PeV	
Zenith angle	$<60°$	
Total number of primaries		
training data	8.0782×10^8 events	3.5971×10^9 events
test data	1.0187×10^8 events	$\sim 1.0 \times 10^9$ events
AS core position	<300 m from the array center	

However, for proton-showers, these heatmaps indicate that the features of multiple regions, both core and periphery, are utilized for separation, irrespective of the shower energy. These results show that the CNN exhibits flexible learning due to the conditions.

4. Application of a CNN with shower density image-like data to a cosmic-gamma-ray observation

The results in section 3 show that the developed CNN method demonstrates better separation performance than other developed methods. In this section, we applied the CNN method to a MC dataset simulating a gamma-ray observation of the Crab Nebula, one of the standard candles for gamma-ray astronomy, in the energy range 10–100 TeV.

4.1. Monte Carlo data of gamma-showers and background hadronic-showers from the Crab Nebula

We generated gamma-showers originating from the Crab Nebula and background hadronic-showers using Monte Carlo simulations with the CORSIKA 7.6400 program and used them as the training and testing data. Table 4 summarizes the simulation settings, where we generated gamma-showers from the Crab Nebula orbit with a power-law energy spectrum index of -2.6 and incident zenith angles smaller than 60 degrees in the energy range from 300 GeV to 10 PeV. The training and test data comprised 8.0782×10^8 and 1.0187×10^8 events, respectively, corresponding to approximately 80 years of observation by the Tibet air shower array. We also generated 3.5971×10^9 events for the training data of background cosmic rays from the Crab Nebula orbit. For the test data of background cosmic rays, we generated events using the following method. Initially, we generated a total of 3.9452×10^9 cosmic-ray events uniformly across all sky and extracted events so that the zenith angle distribution matched that of the Crab Nebula. The number of these events corresponds to approximately 1.0×10^9 events of background cosmic rays coming from the direction of the Crab Nebula. The cosmic-ray chemical composition model was based on the work of Shibata *et al* [37]. The model reproduces the results from the few-teraelectron-volt regions obtained by direct observations to the petaelectron-volt region obtained by the Tibet ASγ experiment.

We calculated the detector response using the Geant4.10.02 software following the procedure described in section 3.1.2 for showers with the incident positions within a radius of 300 m from the center of the Tibet shower array. To select only well-reconstructed shower events, we applied selection conditions 1 to 3 described in section 3.1.3, and the reconstructed zenith angle must be less than 40 degrees. For gamma-ray test data, we further imposed the selection condition that the angular distance between the arrival direction and the gamma-ray source must be less than the on-source window $6.9°/\sqrt{\Sigma\rho/\mathrm{m}^{-2}}$ [1].
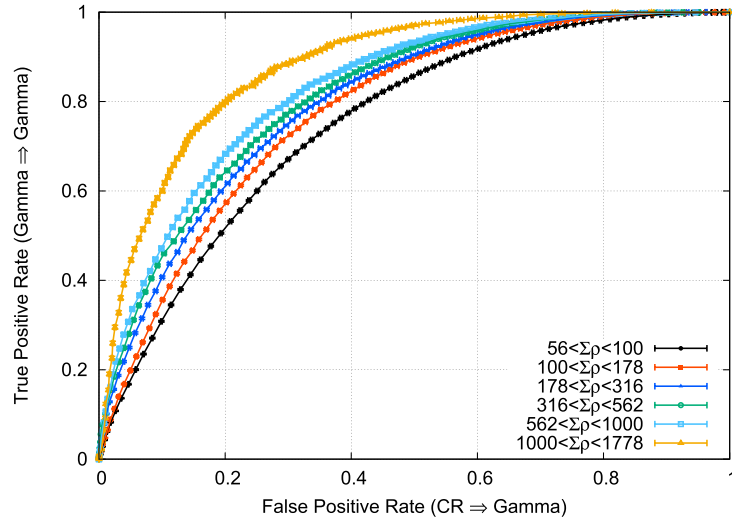
4.1.1. Separation performance

The separation performance is summarized in table 5. The reconstructed showers were divided into six $\Sigma\rho$ bins within the $56\,\mathrm{m}^{-2} < \Sigma\rho < 1778\,\mathrm{m}^{-2}$ range. The second column in table 5 represents the representative energy defined by the logarithmic mean, with a minimum of 11.4 TeV and a maximum of 115 TeV. The ROC curves for each energy bin (figure 15) move toward the upper-left direction as the energy increases, indicating an improvement in the separation performance. The third column in table 5 shows the AUC value for each bin, which is 0.753 at 11.4 TeV and increases as the energy increases, reaching 0.879 at 115 TeV.

4.1.2. Improvement of the detection significance of gamma rays from the Crab Nebula
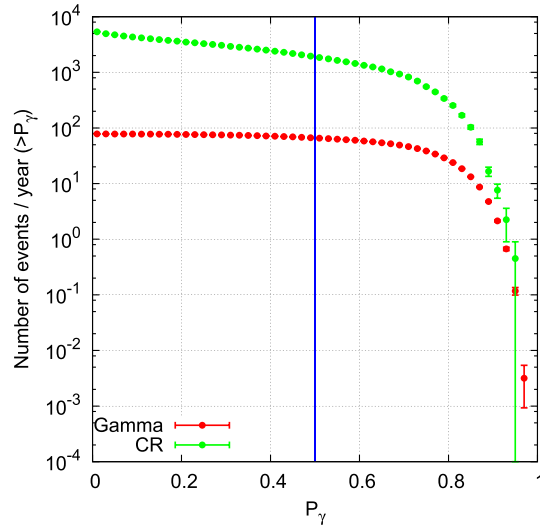
The detection sensitivity of gamma rays depends on the survival ratio of the gamma rays and the background hadron cosmic rays after the separation. Figure 16 shows the change in the number of surviving events versus P_γ values for gamma rays and background cosmic rays in the $316\,\mathrm{m}^{-2} < \Sigma\rho < 562\,\mathrm{m}^{-2}$ (45.0 TeV) bin. The vertical axis represents the expected number of events, converted to a one-year observation period, where the red points indicate the gamma rays and the green points represent background cosmic rays. As P_γ increases,

Table 5. Summary of separation performance.

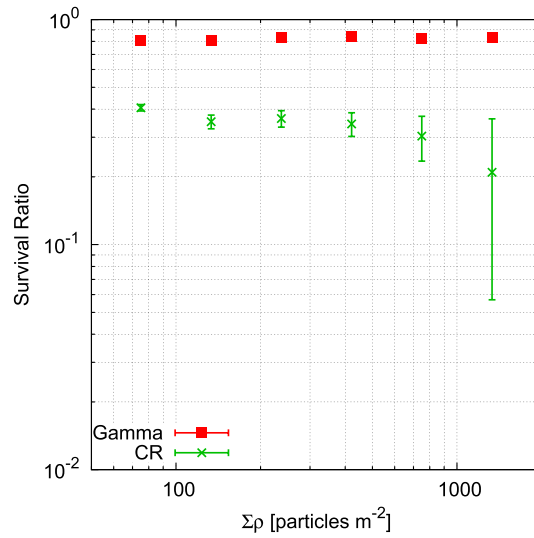| $\Sigma\rho$ (m^{-2})	Energy (TeV)	AUC	Survival ratio		ε
			R_γ	R_{CR}	
56–100	11.4	$0.753^{+0.003}_{-0.003}$	0.809 ± 0.002	0.406 ± 0.013	1.269 ± 0.004
100–178	17.9	$0.781^{+0.003}_{-0.003}$	0.809 ± 0.003	0.352 ± 0.025	1.364 ± 0.005
178–316	28.4	$0.800^{+0.004}_{-0.004}$	0.836 ± 0.003	0.363 ± 0.031	1.386 ± 0.005
316–562	45.0	$0.815^{+0.004}_{-0.004}$	0.839 ± 0.003	0.344 ± 0.041	1.431 ± 0.006
562–1000	71.5	$0.830^{+0.006}_{-0.005}$	0.827 ± 0.004	0.304 ± 0.068	1.501 ± 0.008
1000–1778	115	$0.879^{+0.005}_{-0.005}$	0.834 ± 0.006	0.210 ± 0.152	1.817 ± 0.012

Figure 15. ROC curves for the separation with a CNN. Black shows the results for events in $56\,\mathrm{m}^{-2} < \Sigma\rho < 100\,\mathrm{m}^{-2}$ (11.4 TeV), red shows the result for $100\,\mathrm{m}^{-2} < \Sigma\rho < 178\,\mathrm{m}^{-2}$ (17.9 TeV), blue shows the result for $178\,\mathrm{m}^{-2} < \Sigma\rho < 316\,\mathrm{m}^{-2}$ (28.4 TeV), green shows the result for $316\,\mathrm{m}^{-2} < \Sigma\rho < 562\,\mathrm{m}^{-2}$ (45.0 TeV), cyan shows the result for $562\,\mathrm{m}^{-2} < \Sigma\rho < 1000\,\mathrm{m}^{-2}$ (71.5 TeV), and orange shows the result for $1000\,\mathrm{m}^{-2} < \Sigma\rho < 1778\,\mathrm{m}^{-2}$ (115 TeV).
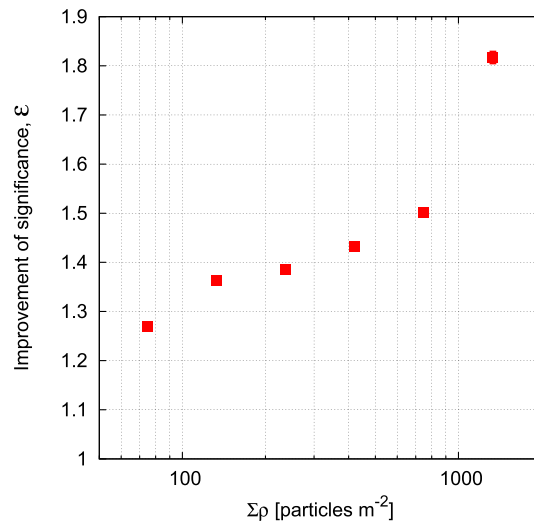

Figure 16. Expected numbers of gamma-showers and hadronic-showers per year observed by the Tibet air shower array. Each plot is the result of the event for $316\,\mathrm{m}^{-2} < \Sigma\rho < 562\,\mathrm{m}^{-2}$ corresponding to an equivalent gamma-ray energy of 45.0 TeV. The red plots indicate gamma-showers, and the green plots show hadronic-showers. The vertical blue line represents cut $P_\gamma = 0.5$.

the number of surviving cosmic-ray events and gamma-ray events decreases. However, the decrease is more significant for the cosmic-ray events. We defined gamma-like events with P_γ greater than a threshold of 0.5. We then calculated the survival ratio of gamma-showers and hadron-showers within each $\Sigma\rho$ bin, denoted

Figure 17. Expected survival ratio of gamma rays (red) and cosmic rays (green) after γ/CR-separation using the optimized cut P_γ.

Figure 18. Improvement of significance, ε, after γ/CR-separation using the optimized cut P_γ.

by R_γ and R_{CR}, respectively. The numerical values of R_γ and R_{CR} are given in columns 4 and 5 of table 5, respectively, and are plotted as a function of $\Sigma\rho$ in figure 17. The survival ratio of gamma rays is approximately 0.8 and remains nearly constant irrespective of energy. On the other hand, the survival ratio of cosmic rays is 0.406 at 11 TeV, decreasing with increasing energy to 0.210 at 115 TeV. When the number of background cosmic rays is sufficiently greater than the number of gamma rays, the significance of gamma-ray excess before and after γ/CR-separation can be approximated as follows: $N_\gamma/\sqrt{N_{\mathrm{CR}}}$ and $N'_\gamma/\sqrt{N'_{\mathrm{CR}}}$, where N_γ and N'_γ are the number of gamma-ray events before and after γ/CR-separation and N_{CR} and N'_{CR} are the number of cosmic-ray events before and after γ/CR-separation, respectively. In figure 16, N_γ and N_{CR} correspond to the number of events at the point on $P_\gamma = 0$, whereas N'_γ and N'_{CR} correspond to those at $P_\gamma = 0.5$. To evaluate the improvement of the significance of gamma-ray excess events, we calculated ε as follows:

$$\varepsilon = \left(\frac{N'_\gamma}{\sqrt{N'_{\mathrm{CR}}}} \right) \Big/ \left(\frac{N_\gamma}{\sqrt{N_{\mathrm{CR}}}} \right) = \frac{R_\gamma}{\sqrt{R_{\mathrm{CR}}}}. \tag{7}$$

Figure 18 and the sixth column of table 5 show the ε values for each $\Sigma\rho$ bin, which increase with increasing energy, from 1.3 at 11.4 TeV to 1.8 at 115 TeV.

5. Summary

This work developed separation methods for gamma-showers and hadronic-showers using the detected particle density measured by the Tibet air shower array. We investigated the characteristics and the separation performance of three methods. To compare the performances, we applied each separation method to air showers generated by simulations of vertically incident 10 TeV gamma rays and 21 TeV protons, and 100 TeV gamma rays and 165 TeV protons. First, we investigated the separation performance using a single feature value among seven, representing the spread of particle density in the air showers and $\Sigma\rho$. The feature value that shows the highest separation performance is R, which has an AUC value of 0.701 for 10 TeV-γ/p-separation and an AUC of 0.808 for 100 TeV-γ/p-separation. Next, we tested an MLP with the seven feature values. The MLP has AUC values of 0.761 for 10 TeV-γ/p-separation and 0.854 for 100 TeV-γ/p-separation. The AUC value of the MLP is approximately 5% greater than the best case of the single-feature-value method, indicating an improvement in separation performance. In addition, the practical feature values vary depending on the primary energy. Next, we tested a CNN with image data converted from the detected particle density data. The CNN has AUC values of 0.781 and 0.901 for 10 TeV- and 100 TeV-γ/p-separations, respectively, which are approximately 5% greater than the corresponding AUC values of the MLP. From the heatmaps generated using Grad-CAM, we found that the parts of the particle spread in air showers that significantly affect the separation vary depending on the energy. The CNN method, which has the advantage of automatically optimizing filter parameters for the feature extraction, demonstrates better separation performance than the other two methods.

To investigate the effectiveness of the CNN method in teraelectron volt gamma-ray observations, we separated observed air showers in the energy range from tens of TeV to 100 TeV using a MC dataset simulating the Tibet air shower array. When a CNN is applied to γ/CR-separation using simulation data from the Crab Nebula, the significance of gamma-ray excess events is improved compared with the case without separation. The sensitivity is enhanced by a factor of 1.3 at 11.4 TeV and 1.8 at 115 TeV.

These results show the effectiveness of the CNN method for gamma-ray observations with the Tibet air shower array. The method developed in the present study can be helpful in the reanalysis of gamma-ray source searches reported before construction of the MD. Alternatively, other groups have explored the application of a CNN for determining the primary energy and arrival direction [38]. Therefore, the CNN analysis method developed in the present study can conceivably be applied to primary particle reconstruction in the data analysis of the Tibet air shower array.

Data availability statement

The data cannot be made publicly available upon publication because no suitable repository exists for hosting data in this field of study. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

ORCID iDs

Sousuke Okukawa ⬡ https://orcid.org/0009-0002-0664-2031
Kinya Hibino ⬡ https://orcid.org/0000-0001-9259-6371
Yusaku Katayose ⬡ https://orcid.org/0000-0003-4836-7621
Takashi Sako ⬡ https://orcid.org/0000-0002-2589-8561
Atsushi Shiomi ⬡ https://orcid.org/0000-0002-2616-0059

References

[1] Amenomori M *et al* 2019 *Phys. Rev. Lett.* **123** 051101
[2] Abeysekara A U *et al* 2017 *Science* **358** 911–4
[3] Cao Z *et al* 2021 *Nature* **594** 33–36
[4] Gaisser T L 1990 *Cosmic Rays and Particle Physics* (Cambridge University Press)

[5] Sako T K, Kawata K, Ohnishi M, Shiomi A, Takita M and Tsuchiya H 2009 *Astropart. Phys.* **32** 177–84

[6] Amenomori M *et al* 2015 *Astrophys. J.* **813** 98

[7] Abeysekara A U *et al* 2017 *Astrophys. J.* **843** 39

[8] DeMitri I and Salamida F 2004 *Nucl. Instrum. Methods Phys. Res.* A **525** 132–6

[9] Amenomori M *et al* 2003 *Astrophys. J.* **598** 242–9

[10] Abdo A A *et al* 2010 *Science* **331** 739–42

[11] Okukawa S *et al* 2023 *Proc. Sci.* **444** 786

[12] Erdmann M, Glombitza J and Walz D 2017 *Astropart. Phys.* **97** 46–53

[13] Ivanov D *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 015006

[14] Amenomori M *et al* 1992 *Phys. Rev. Lett.* **69** 2468–71

[15] Amenomori M *et al* 2021 *Phys. Rev. Lett.* **126** 141101

[16] Amenomori M *et al* 2009 *Astrophys. J.* **692** 61–72

[17] Amenomori M *et al* 2006 *Phys. Lett.* B **632** 58–64

[18] Amenomori M *et al* 2008 *Nucl. Phys.* B **175–176** 318–21

[19] Amenomori M *et al* 2009 *J. Phys. Soc. Japan* A **78** 206–9

[20] Amenomori M *et al* 2011 *Adv. Space Res.* **47** 629–39

[21] Bishop C and Nasrabadi N M 2006 *Pattern Recognition and Machine Learning* (Springer) p 7

[22] Heck D, Knapp J, Capdevielle J N, Schatz G and Thouw T 1998 CORSIKA: a Monte Carlo code to simulate extensive air showers *Report No.* FZKA-6019 (Forschungszentrum Karlsruhe)

[23] Amenomori M *et al* 2005 *Astrophys. J.* **633** 1005–12

[24] Pierog T, Karpenko I, Katzy J M, Yatsenko E and Werner K 2015 *Phys. Rev.* C **92** 034906

[25] Battistoni G *et al* 2015 *Ann. Nucl. Energy* **82** 10–18

[26] Amenomori M *et al* 2008 *Astrophys. J.* **678** 1165–79

[27] Kato S *et al* 2021 *Exp. Astron.* **52** 85–107

[28] Amenomori M *et al* 1990 *Nucl. Instrum. Methods Phys. Res.* A **288** 619–31

[29] Amenomori M *et al* 2017 *Proc. Sci.* **1001** 484

[30] Amenomori M *et al* 2006 *J. Phys.: Conf. Ser.* **47** 51–58

[31] Fawcett T 2006 *Pattern Recognit. Lett.* **27** 861–74

[32] Chollet F *et al* 2015 Keras (GitHub) (available at: https://github.com/fchollet/keras)

[33] Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z 2016 *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* pp 2818–26

[34] Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R 2014 *J. Mach. Learn. Res.* **15** 1929–58 (available at: https://jmlr.org/papers/v15/srivastava14a.html)

[35] Lin M, Chen Q and Yan S 2014 Network in network (arXiv:1312.4400)

[36] Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D and Batra D 2017 *2017 IEEE Int. Conf. on Computer Vision (ICCV)* pp 618–26

[37] Shibata M, Katayose Y, Huang J and Chen D 2010 *Astrophys. J.* **716** 1076–83

[38] Kalashev O 2020 *J. Phys.: Conf. Ser.* **1525** 012001