Development of convolutional neural networks for an electron-tracking Compton camera

Ikeda, Tomonori; Takada, Atsushi; Abe, Mitsuru; Yoshikawa, Kei; Tsuda, Masaya; Ogio, Shingo; Sonoda, Shinya; Mizumura, Yoshitaka; Yoshida, Yura; Tanimori, Toru

Ikeda, Tomonori ...[et al]. Development of convolutional neural networks for an electron-tracking Compton camera. Progress of Theoretical and Experimental Physics 2021, 2021(8): 083F01.

PTEP
Prog. Theor. Exp. Phys. **2021**, 083F01 (13 pages)
DOI: 10.1093/ptep/ptab091

Development of convolutional neural networks for an electron-tracking Compton camera

Tomonori Ikeda[1,*], Atsushi Takada[1], Mitsuru Abe[1], Kei Yoshikawa[1], Masaya Tsuda[1], Shingo Ogio[1], Shinya Sonoda[1], Yoshitaka Mizumura[2], Yura Yoshida[1], and Toru Tanimori[1]
[1]*Graduate School of Science, Kyoto University, Kitashirakawa Oiwakecho, Sakyo, Kyoto, Kyoto, 606-8502, Japan*
[2]*Scientific Balloons, Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, Yoshinodai 3-1-1, Chuou, Sagamihara, Kanagawa, 252-5210, Japan*
*E-mail: ikeda.tomonori.2s@kyoto-u.ac.jp

Received May 14, 2021; Revised July 2, 2021; Accepted July 5, 2021; Published July 14, 2021

The Electron-Tracking Compton Camera (ETCC), which is a complete Compton camera that tracks Compton scattering electrons with a gas micro time projection chamber, is expected to open up MeV gamma-ray astronomy. The technical challenge for achieving several degrees of the point-spread function is precise determination of the electron recoil direction and the scattering position from track images. We attempted to reconstruct these parameters using convolutional neural networks. Two network models were designed to predict the recoil direction and the scattering position. These models marked 41° of angular resolution and 2.1 mm of position resolution for 75 keV electron simulation data in argon-based gas at 2 atm pressure. In addition, the point-spread function of the ETCC was improved to 15° from 22° for experimental data from a 662 keV gamma-ray source. The performance greatly surpassed that using traditional analysis.

Subject Index F10

1. Introduction

Gamma-ray observations in the low-energy gamma-ray band from a hundred keV to several MeV address various astrophysical phenomena: nucleosynthesis and the explosion mechanism in supernovae [1–3], the particle acceleration process in active galactic nuclei or gamma-ray bursts [4–7], and the electron–positron annihilation line in the Galactic center region [8–10]. However, observation in this energy band remains poorly explored compared to the X-ray, GeV, and TeV bands, known as the “MeV gap.” Even now, COMPTEL [11], onboard the Compton Gamma Ray Observatory launched in 1991, is the most sensitive observer of the MeV sky. The causes of this stagnation in MeV observation are the huge gamma-ray backgrounds from the atmosphere (albedo) and generated in the instruments by cosmic-ray interactions. In addition, conventional Compton cameras like COMPTEL have unclean images according to the Compton circle due to the lack of direction information for the recoil electron. Such pseudo imaging is not capable of background rejection and does not provide quantitative imaging [12]. Thus, it was pointed out that conventional Compton cameras need the additional parameters of the recoil direction of the Compton electron in order to reduce backgrounds [13].

The Electron-Tracking Compton Camera (ETCC) is a complete Compton camera that tracks recoil electrons. It can record all the information on Compton kinematics to overcome the problem of conventional Compton cameras. Hence, ETCC determines a unique incident gamma-ray direction [12]. A key device of the ETCC is the tracking detector that detects the recoil electron, since the

Funded by SCOAP^3

point-spread function (PSF) is highly dependent on the determination accuracy of the recoil direction and the scattering position [14]. Thus, the SMILE group (Sub-MeV gamma-ray Imaging Loaded-on-ballon Experiment) developed the gas time projection chamber (TPC) based on μ-PIC [15], which is one of the micro-pattern gas detectors. Although the gas TPC gives us precise track information for low-energy electrons, these electrons are influenced by multiple scattering and make complex track images. The technical challenge is the development of the analysis method with an exact determination of the scattering position and the recoil direction from these track images.

Feature extraction from images is the field of computer vision. Convolutional neural networks (CNNs) have achieved great success in image classification. Since 2010, many architectures like AlexNet [16], VGG [17], GoogLeNet [18], and ResNet [19] were proposed in the ImageNet Large Scale Visual Recognition Challenge [20] and achieved dramatic progress. In particle physics and astrophysics experiments, modern machine learning techniques have been actively applied and developed. They perform several tasks, such as signal–background recognition, particle identification, and event reconstruction. Machine learning has shown significant performance improvements over the traditional way based on domain knowledge. In the gas TPC application, a CNN was developed for polarization extraction from photoelectron track images taken with X-ray polarimeters and improved the polarization sensitivity by 10%–20% [21]. The NEXT experiment, searching for neutrinoless double-beta decay, utilized a CNN to identify electron–positron pair production from the topological signature; the signal efficiency was improved compared to non-CNN-based analysis [22]. In our case, a CNN would also be a promising approach.

In this paper we describe the design of CNNs to predict the scattering position and the electron recoil direction from track images taken by the TPC. In addition, we evaluate the imaging performance for simulation and experimental data, and compare with a traditional method.

2. The Electron-Tracking Compton Camera

The dominant interaction process of MeV gamma-rays and materials is Compton scattering. A schematic view of the Compton kinematics is shown in Fig. 1. A conventional Compton camera obtains the scattering gamma-ray energy E_γ, the absorption position of the scattering gamma-ray $P_{\rm abs}$, the scattering position of the gamma-ray $P_{\rm sct}$, and the electron recoil energy $K_{\rm e}$ as well as the electron recoil direction $\vec{e}$. Then, the scattering angle ϕ is written as

$$\cos\phi = 1 - \frac{m_{\rm e}c^2}{E_\gamma + K_e}\frac{K_{\rm e}}{E_\gamma}. \quad (1)$$

Conventional Compton cameras are insufficient to resolve the kinematic equation of Compton scattering due to the lack of the electron recoil direction. While the incident gamma direction $\vec{s}$ is only reconstructed as a Compton circle, the overlapping of many Compton circles determines the source location. In a conventional Compton camera, the PSF is only defined by the angular resolution measure (ARM). The ARM means the angular distance from the reconstructed Compton circle to the known source location or the determination accuracy of the Compton scattering angle ϕ. The ARM depends on the determination accuracies of the absorption point and the scattering point of a gamma-ray, and the energy resolution.

COMPEL was incredibly successful via Compton imaging techniques. However, the achieved sensitivity was modest. One of the most important things we learned from COMPTEL is that huge gamma-ray backgrounds are generated in the satellite itself by cosmic rays. Under such backgrounds, the sensitivity is overestimated if we only use the ARM. Schönfelder reported the background in

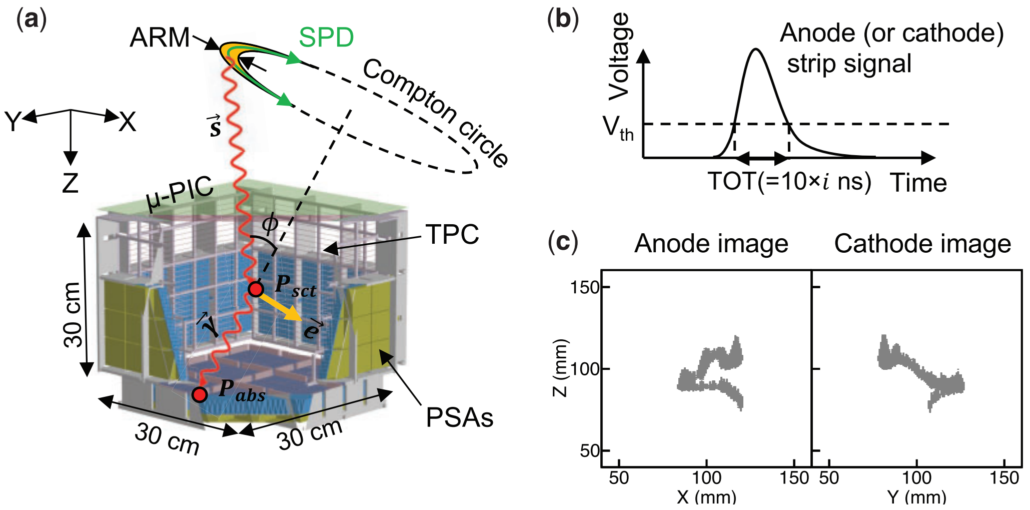

Fig. 1. (a) Schematic view of the ETCC on the SMILE-2+ experiment. The incident gamma-ray coming from the source location $\vec{s}$ undergoes Compton scattering into the TPC filled with argon-based gas. The energy, scattering position P_{sct}, and direction $\vec{e}$ of the recoil electron are reconstructed by the TPC based on μ-PIC. Finally, the scattering gamma-ray $\vec{\gamma}$ is absorbed at P_{abs} by pixel scintillator arrays. (b) Schematic image of the time over threshold. (c) Track images of recoil electrons with 160 keV from μ-PIC anode and cathode strips.

COMPTEL and concluded that the most sensitive way to increase the sensitivity would be to reduce the instrumental background rate [13]. He also proposed the measurement of recoil electrons in order to suppress the background. From this effort, we developed ETCC as a next-generation Compton telescope.

The ETCC provides all the parameters of Compton scattering, including the electron recoil direction, with the tracking detector. Figure 1(a) shows a schematic view of the ETCC on the SMILE-2+ experiment, which is our second balloon experiment to observe celestial objects at high altitudes [23]. The ETCC completely resolves the Compton kinematics and determines the unique incident gamma-ray direction $\vec{s}$ using the following equation:

$$\vec{s} = \frac{E_\gamma}{E_\gamma + K_\mathrm{e}}\vec{\gamma} + \frac{\sqrt{K_\mathrm{e}(K_\mathrm{e} + 2m_\mathrm{e}c^2)}}{E_\gamma + K}\vec{e}. \tag{2}$$

The ETCC gives the complete PSF from the ARM and the scattering plane deviation (SPD) [12], which is the accuracy of the scattering plane determination. The ETCC can achieve high background rejection power by localizing the arrival direction on the Compton circle thanks to the SPD.

The ETCC of the SMILE-2+ experiment consists of pixel scintillator arrays (PSAs) and a gas TPC based on a micro pattern gas detector, μ-PIC [15]. The 36 PSAs with a thickness of 26 mm and the 18 × 4 PSAs with a thickness of 13 mm are arranged below and to the sides of the TPC, respectively. Each PSA is made of GSO (Gd_2SiO_5:Ce) scintillators of 8 × 8 pixels with a pixel size of $6 \times 6\,\mathrm{mm}^2$. PSAs are deployed as absorbers to detect the energy E_γ and the position P_{abs} of the scattering gamma-ray. The TPC is filled with argon-based gas (95% Ar + 3% CF_4 + 2% C_4H_{10} at a pressure of 2 atm) as a Compton-scattering target and has a drift length of 30 cm. The electron track and energy information was detected by μ-PIC, which has 768 × 768 strips with a pitch of 400 μm; the detection volume is $30 \times 30 \times 30\,\mathrm{cm}^3$. In order to reduce the power consumption, every two strips of μ-PIC are grouped. Thus, the readout strip pitch is 800 μm in the SMILE-2+ experiment.

The anode and cathode strips of μ-PIC are linearly arranged in the X and Y directions, respectively. The Z direction is reconstructed from the measured drift velocity and drift time. Thus, two-dimensional images (XZ and YZ) are obtained. The strip signals are processed in FE2009bal CMOS ASIC chips [24] including preamplifiers, shapers, and comparators. Each amplified signal is compared to the threshold voltage and synchronized with 100 MHz (10 ns) clocks. Finally, the addresses of μ-PIC strips with the time-over-threshold (TOT) are recorded. Figure 1(b) shows the schematic image of the TOT. The TOT is the time between rising and falling edges crossing the threshold voltage, which is correlated with the convolution of the charge information and the track length along the Z-axis on each strip. Since the TOT is measured by the comparator at 100 MHz, it is digitized at 10 ns. Figure 1(c) shows a typical electron track obtained by μ-PIC. The lengths of the gray pixels along the Z-axis on each strip correspond to the TOT multiplied by the drift velocity. The low-density scatter medium and the fine-interval read-out detector of μ-PIC provide fine track images. The scattering position $P_{\rm sct}$ and electron recoil direction $\vec{e}$ are reconstructed from these two images.

The traditional method to determine the scattering position is to utilize the skewness of the TOTs, which relies on the fact that the stopping power of recoil electrons depends on their residual energy, and the end-point of the track has large TOTs. The skewness is written by

$$S = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\langle (x - \langle x \rangle)^3 \rangle}{\langle (x - \langle x \rangle)^2 \rangle^{3/2}},$$

and is represented as the dimensionless ratio between the third and second moments of the TOTs. The skewness provides information on whether the endpoint or the start point is on the right or left side of the image. Consequently, the maximum or minimum of a track image is adopted as a scattering position. After calculating the scattering position, the recoil direction is determined by a linear fitting algorithm using TOTs within 4 mm of the reconstructed scattering point to prevent multiple scattering effects. This method is suitable for relatively straight tracks, while it is less accurate for curved tracks from multiple scattering.

3. Convolutional neural networks

A CNN is an example of an artificial neural network. It has an input layer, hidden layers, and an output layer. A hidden layer is generally composed of a combination of a convolutional layer, a pooling layer, and a fully connected layer. The convolutional layer, which is the most important element in the CNN, automatically extracts features from the image by a weighting matrix, called a filter or kernel. The pooling layer is used to down-sample the feature maps extracted in the convolutional layer by taking an average or maximum. In a fully connected layer, all nodes are connected to all activations in the previous layer. A deep hierarchical structure of the convolutional and pooling layers has expressive capacity; a shallow layer learns low-level features or local correlations, and a deep layer learns high-level features or global correlations. Thus, general CNNs are made up of stacks of convolutional and pooling layers.

We implemented two independent network models to predict the scattering position and the electron recoil direction using the Keras/Tensorflow framework [25]. The common point in both network models is that the input layer takes two-channel images of 256 × 512 pixels, corresponding to $204.8 \times 194.56\ \mathrm{mm}^2$, in the XZ and YZ dimensions. Thus, the number of input parameters (ϕ, $\cos\theta$) is $2 \times 256 \times 512$ $(= 262144)$. In this section we describe the characteristics of these network models.

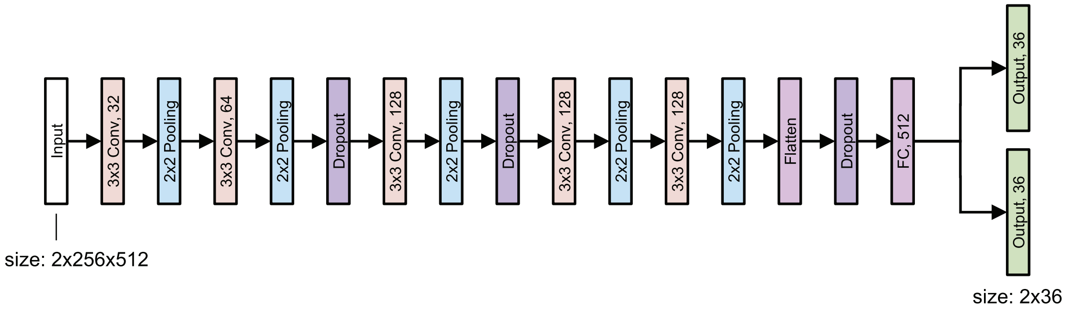

Fig. 2. Network model to predict the electron recoil direction. The white and green boxes are the input and output layers, respectively. The orange boxes represent convolution layers with stride 1, which are denoted as (filter size) Conv, (number of filters). The blue boxes are maximum pooling layers with a filter size of 2 × 2 and stride 2. The purple boxes are dropout layers with the rate set to 20%. The pink boxes represent fully connected or flattening layers.

3.1. Network model to predict the electron recoil direction

In 2014, the Visual Geometry Group (VGG) model [17] received high praise in the classification category of the image recognition competition. This architecture simply increases the number of layers by connecting the convolution layer and the fully connected layer. The reduction of the filter size enables the implementation of a deeper network giving the ability to learn high-level features. We have designed a network model based on the VGG model to predict the electron recoil direction.

A schematic view of the network model is shown in Fig. 2. This model has two output layers to predict the cosine of zenith angle $\cos\theta$ and the azimuth angle ϕ of the electron recoil direction. The prediction of each angle is divided into 36 classes, each having a width of 0.055 and 10 degrees, respectively. Thus, the number of output parameters is 2×36. The output layers are activated by softmax functions to compute each of the probability distributions. The hidden layer is a hierarchical representation of convolutional and pooling layers. All convolutional layers have a filter size of 3×3 and stride 1. All pooling layers take the maximum value with a filter size of 2×2 and stride 2. In addition, three dropout layers, in which some number of layer outputs are randomly ignored, are deployed to prevent over-fitting. In the final stage, two-dimensional feature maps are converted to a one-dimensional vector in a flattening layer and connected to the fully connected layer with 512 nodes. All layers except for pooling and output layers are activated by the rectified linear unit (ReLU) [26].

3.2. Network model to predict the scattering position

The accuracy of the scattering position is expected to be the pixel size resolution. While the VGG model is able to learn high-level features with the deep layer, the detailed position information is lost at the same time as the size of the feature maps is reduced. For spatially dense predictions like semantic segmentation, which is the process of separating an image into some object of interest at a pixel level, fully convolutional networks (FCNs) [27] and U-Net [28] were developed. These architectures have an encoder part followed by a decoder part. The encoder part is a typical classifier network like the VGG. The decoder part expands the feature map and projects onto high-resolution images. In order to predict the scattering position, we built the network model on the U-Net architecture.

A schematic view of the network model is depicted in Fig. 3. In order to predict the scattering positions (x, y, z), this model has three output layers, which have linear activation functions. The size of the output parameter is 3×1. The scattering position is predicted on a regression problem. The

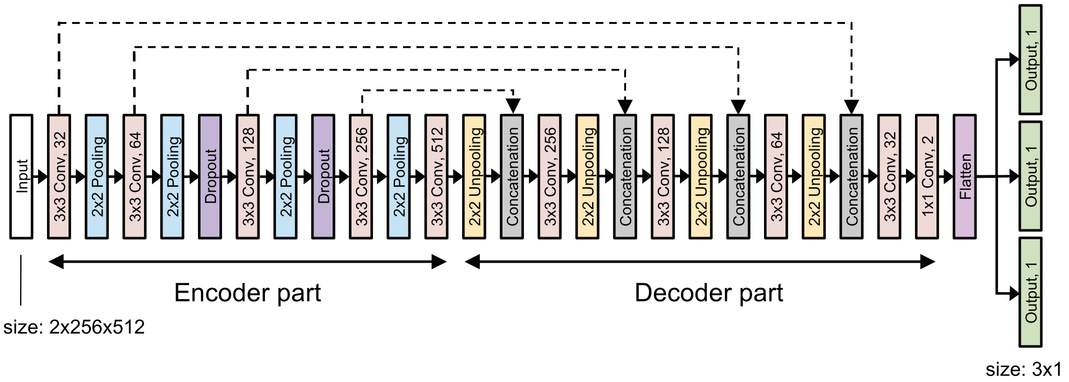

Fig. 3. Network model to predict the scattering position. The white and green boxes are the input and output layers, respectively. The orange boxes represent convolution layers with stride 1, which are denoted as (filter size) Conv, (number of filters). The blue boxes are maximum pooling layers with a filter size of 2 × 2 and stride 2. The purple boxes are dropout layers with the rate set to 20%. The yellow boxes are the unpooling layers. The pink box represents a flatten layer. The black dotted arrows indicate the concatenation operation to combine the output of the convolution layers in the encoder part with the decoder part.

hidden layer is comprised of the encoder and decoder part. The first encoder section extracts features at different scale sizes with down-sampling as well as the network to predict the electron recoil direction. The convolutional layers have 3 × 3 kernels with stride 1, and the pooling layers are the maxpooling operation and stride 2. In the second decoder section, the pooling layers are replaced with unpooling layers for up-sampling and expanding feature maps. We deploy the convolution transpose layer as an unpooling operation with a filter size of 2 × 2 and stride 2. In addition, up-sampling images are connected by down-sampling images in concatenation layers and recover the high spatial resolution. The final up-sampling layer is flattened and connected to the output layers. All the layers except for the pooling and output layers are activated by the ReLU.

4. Training and validating the network via a Monte Carlo simulation

4.1. Preparing a training and validation data set

The training and validation data are prepared by the Monte Carlo (MC) simulation package of the SMILE experiment based on Geant4 [29]. In order to gernerate experimental images, the following procedures are undertaken:

- The full detector geometry is constructed in Geant4. Electron beams are generated in the TPC volume with uniform position and direction. Then, the interaction points and the deposit energies are recorded.
- From the deposit energies, the number of ionized electrons is calculated considering the Fano factor. In addition, the collected positions of ionized electrons on μ-PIC strips are obtained according to the diffusion effect simulated by MAGBOLTZ [30].
- Waveforms of μ-PIC strips are represented by overlapping waveform templates, which are calculated from the simulated μ-PIC pulse signal using Garfiled++ [31]. These waveforms are digitized and TOTs are obtained by comparing to the threshold. Finally, the calculated TOTs are encoded to the experimental data format.

We analyze simulated data and make *XZ* and *YZ* images of 256 × 512 pixels so that the tracks are in the center of the images, like Fig. 1(c). Overall, 630k and 70k MC events for 5–200 keV, which were fully contained in the TPC volume, were generated for the training and validation data sets, respectively. Also, each class has almost the same number of samples; one class has about 17500 samples for the training data set.

4.2. Training and validation network

We trained the networks and optimized the learnable parameters by minimizing the loss function between the ground truth and the prediction. In training the model to predict the electron recoil direction, the ground truth is expressed by a "one-hot" vector $\vec{p}$. For example, the ith class constitutes angles ϕ ranging from $\pi i/18-\pi$ to $\pi(i+1)/18-\pi$. Then, the ith label is 1 and the rest 0 in the one-hot vector. We employed the cross-entropy loss, which is a typical loss function in the classification task, represented by $H(p,q)=-\Sigma_i p_i \log q_i$, where p_i is the ith label of the ground truth and q_i is the predicted one. On the other hand, we employed the mean square root as a loss function to train the model to predict the scattering position. In both cases, an Adam optimizer was used to minimize the loss function with a learning rate of 0.0001. A batch size of four was chosen. We trained the models for ten epochs where validation losses saturated. The huge amount of calculations wer performed on NVIDIA Geforce GTX 2070 super graphics processing units (GPUs).

Figure 4 shows confusion matrices at the ϕ and $\cos\theta$ class predictions of several energies. In 5–10 keV of ϕ, the CNN almost misclassifies and outputs two indexes in any true ϕ. We confirmed that such a high bias prediction appears when the highest predicted probability is very low. Therefore, the angle ϕ is determined by random uniformly in such a case. As the highest predicted probability exceeds a certain threshold, the predicted angle ϕ is calculated as follows:

$$\phi=\mathrm{atan2}\left(\sum_{i=0}^{N_c-1} p_i \sin\phi_i, \sum_{i=0}^{N_c-1} p_i \cos\phi_i\right), \tag{3}$$

where p_i is the predicted probability in the ith class, ϕ_i is the center angle of the ith class, and N_c is the number of classes. The correlation matrix between the calculated angle ϕ and the true angle ϕ are shown in Fig. 5. The high bias prediction in 5–10 keV was reduced due to the uniform random. On the other hand, some bias prediction or nonuniform angular response remains in 10–200 keV. In order to remove the nonuniform response, adding the nonuniform penalty term to the loss function is valuable as discussed in Ref. [21]. We will deploy such an advanced method in our future work. The angle θ is calculated the same way. The confusion matrices and correlation matrices of $\cos\theta$ are described in the bottom panels of Figs. 4 and 5, respectively.

Two values are evaluated. The first is the angular error, $\cos\theta_{\mathrm{err}}=\vec{V}_{\mathrm{t}}\cdot\vec{V}_{\mathrm{p}}$, where $\vec{V}_{\mathrm{t}}$ and $\vec{V}_{\mathrm{p}}$ are the true and predicted recoil vectors, respectively. The second value is the position error, $R_{\mathrm{err}}=|\vec{P}_{\mathrm{t}}-\vec{P}_{\mathrm{p}}|$, where $\vec{P}_{\mathrm{t}}$ and $\vec{P}_{\mathrm{p}}$ are the true and predicted scattering positions, respectively. We obtained these distributions for 70k validation data values. The angular and position resolutions are defined by the 50% area values. In the traditional method, these values are calculated in the same way.

Several reconstructed examples are depicted in Fig. 6. Using the traditional method, high-energy recoil events like the bottom figures could not be reconstructed with high precision due to the multiple scattering effect. On the other hand, the CNN succeeded in resolving the scattering point and recoil direction for such events.

Figures 7(a) and (b) show the energy dependence of the angular resolution and the position resolution compared to the traditional method. The CNN method is better than the traditional method

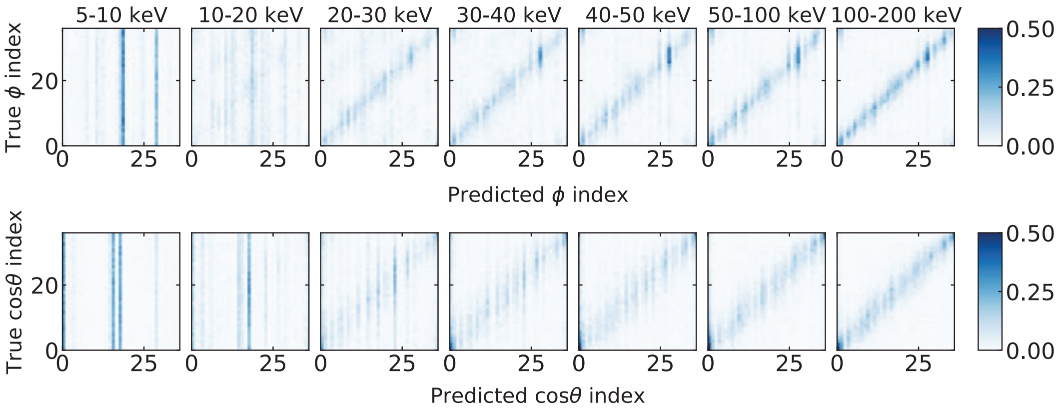

Fig. 4. Confusion matrices at the ϕ (top panel) and $\cos\theta$ (bottom panel) class predictions.

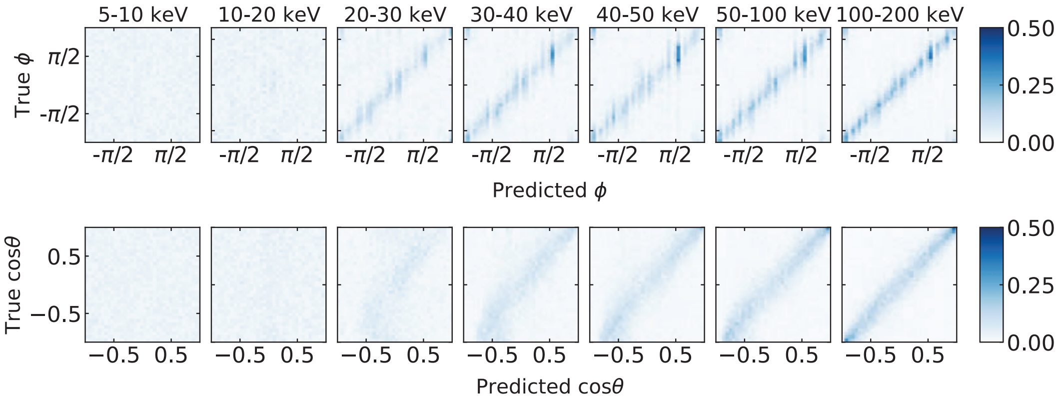

Fig. 5. Top: The correlation matrices between the calculated angle ϕ and the true angle ϕ. Bottom: The correlation matrices between the calculated $\cos\theta$ and the true $\cos\theta$.

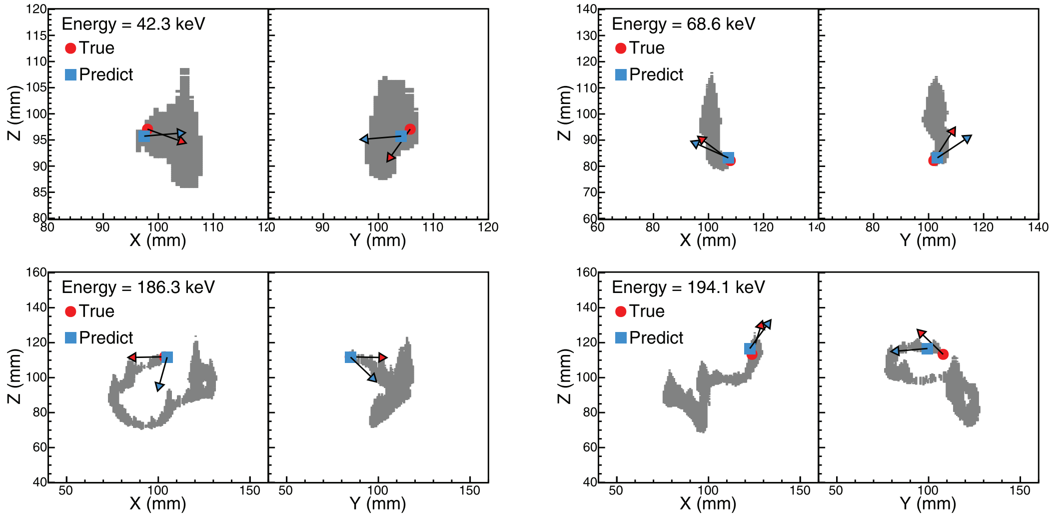

Fig. 6. Four track images with the predicted scattering positions and recoil directions. The red points and arrows are the true scattering points and recoil directions, while the blue squares and arrows are the predictions using the CNN.

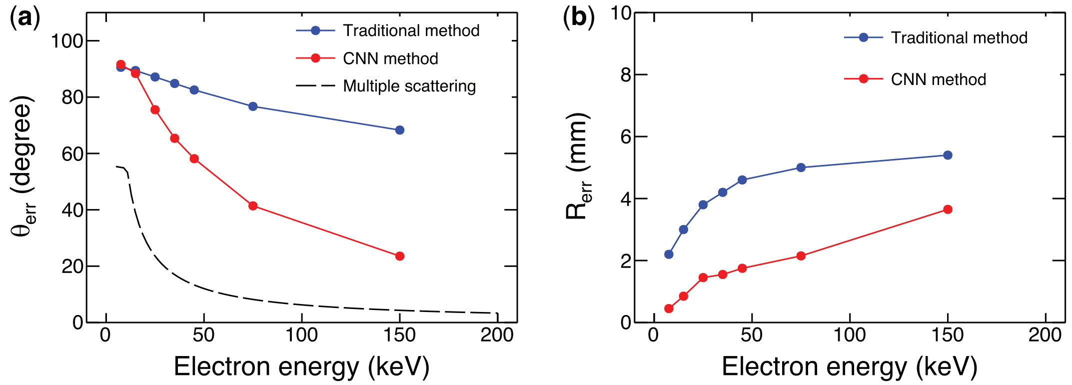

Fig. 7. (a) Angular resolution of the CNN method compared to the traditional method using the simulation data. The red and blue points indicate the CNN and traditional methods, respectively. The black dashed line is the principle limit of multiple scatterings with a 0.8 mm pitch readout. (b) Position resolution of the CNN method compared to the traditional method. The red and blue points are the CNN and traditional methods, respectively.

above 20 keV. In particular, it achieved excellent performance above 50 keV. The angular resolution for 75 keV is 41°. For predicting the scattering position, the CNN method indicates better performance at all energies. The position resolution at 75 keV is 2.1 mm.

5. Performance of ETCC using the CNN method

We conducted the ETCC calibration measurement for the SMILE-2+ instrument on the ground. In order to investigate several energy responses, calibration data for ^{133}Ba (mainly 356 keV), ^{22}Na (511 keV), ^{137}Cs (662 keV), and ^{60}Co (1173 keV and 1332 keV) were taken. In addition, the ^{137}Cs source was placed at several zenith positions separated by 183 cm from the ETCC center. We reanalyzed these calibration data with the CNN method and evaluated the performance of the experimental ETCC data.

5.1. Event selection

The calibration measurements involve not only direct incident gamma-rays from the source but also backgrounds of ambient gamma-rays and cosmic rays. Therefore, we adopted the following event selection to extract direct incident gamma-rays from the sources.

- One hit event: We require one track in the TPC and one hit pixel signal in the PSA, which removed most accidental coincidence events.
- Fiducial volume: We require that an electron track is contained within the fiducial volume of $25.72 \times 25.72 \times 29.8$ cm^3. At the boundary of the detection area, the electric field is distorted due to the supply voltage of the PMT of the PSA around the TPC. Such events which do not retain the original track information are eliminated.
- Fully contained electron: The dE/dx distribution or the correlation between the track length and the energy deposit of the track in the TPC distinguishes the fully contained electrons, escaped electrons from the TPC, and cosmic-ray muons. The dE/dx of a fully contained electron is empirically proportional to $E^{1.72}$ [32] of the recoil electron energy. We determined the fully contained electron band by fitting with a Gaussian function every energy bin with a width of

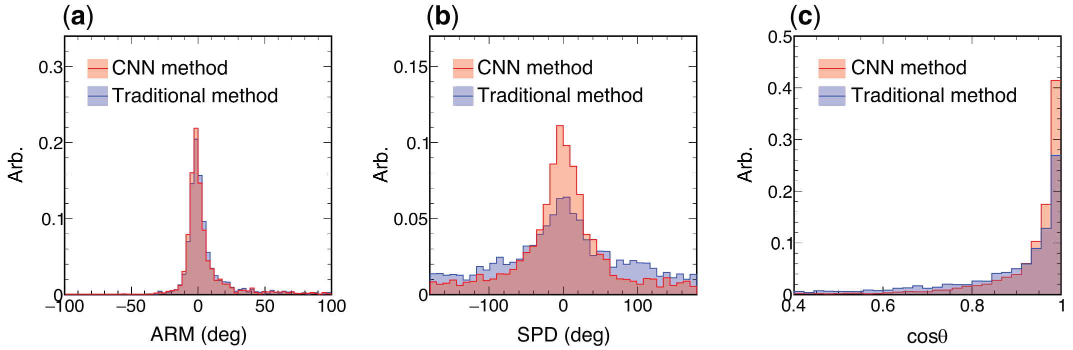

Fig. 8. (a) ARM, (b) SPD, and (c) $\cos\theta$ distributions of the experiment data. The red and blue histograms, which are normalized by the total areas, indicate the CNN and traditional methods, respectively.

10 keV, and the region within the $\pm 3\,\sigma$ quantiles was chosen. More details can be found in Ref [14].

- Compton kinematic test: The angle between the recoil electron direction and the scattering gamma direction is defined as the Compton α angle and calculated geometrically by

$$\cos\alpha_{\rm geo} = \vec{g}\cdot\vec{e}. \tag{4}$$

In addition, this angle is derived by the Compton scattering kinematics:

$$\cos\alpha_{\rm kin} = \left(1 - \frac{m_e c^2}{E_\gamma}\right)\sqrt{\frac{K_e}{K_e + 2m_e c^2}}. \tag{5}$$

Therefore, $\Delta\cos\alpha$ ($= \cos\alpha_{\rm geo} - \cos\alpha_{\rm kin}$) being nearly zero ensures that the reconstructed events are true Compton scattering events. We applied the limit $|\Delta\cos\alpha| < 0.5$ in order to extract Compton scattering events.

- Energy selection: Events with a reconstructed energy within twice the full width at half maximum (FWHM) for the source energy are used to extract the direct components.

After the event selection, the effective areas of the experimental and simulation data for 662 keV were 2.3×10^{-1} cm^2 and 2.1×10^{-1} cm^2, respectively.

5.2. *Imaging performance*

The incident gamma-ray direction is reconstructed by Eq. (2). To evaluate the reconstruction accuracy, we defined the error angle $\cos\theta$ between the reconstructed vector and the source vector. The PSF is the angular resolution of θ. Thus, we calculated the value of half the area of the $\cos\theta$ distribution as the PSF. Figure 8 shows the ARM, SPD, and $\cos\theta$ distribution using the CNN and traditional methods with the ^{137}Cs calibration source at the zenith = 0. The ARM using the CNN method is 10° and is not different from the traditional method. We confirmed that the supply voltage of the PSAs around TPC distorted the electric field, and this dominated the uncertainty of the scattering position. On the other hand, the SPD was highly improved thanks to reducing the angular resolution of the recoil direction. We obtained 59° (FWHM) for the SPD by fitting a Gaussian function. This benefit gave a sharp $\cos\theta$ distribution and improved the PSF to 15°.

Figures 9(a) and (b) show the dependencies of the PSF on the gamma-ray energy and the incident zenith angle, respectively. For every gamma-ray energy and incident zenith angle, the CNN method performed markedly better than the traditional method.

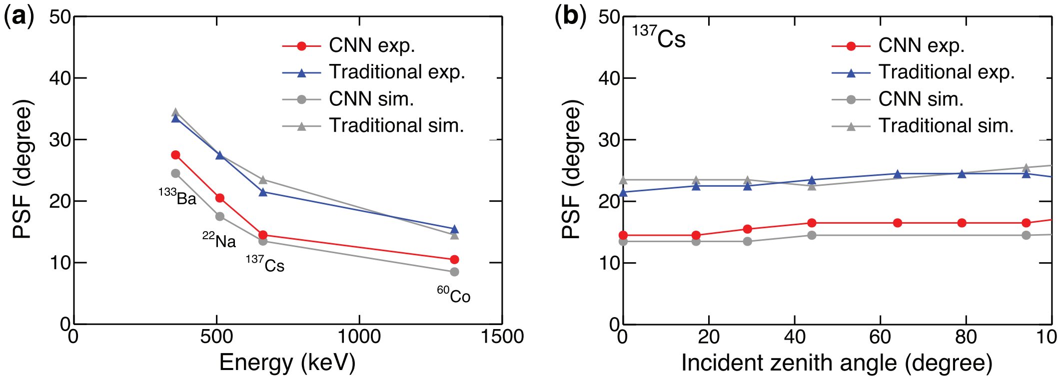

Fig. 9. (a) Gamma-ray energy dependence and (b) incident zenith angle dependence of PSF. The red and blue lines indicate the experimental data using the CNN method and traditional methods, respectively. The gray dots and triangles are the simulation data using the CNN and traditional methods, respectively.

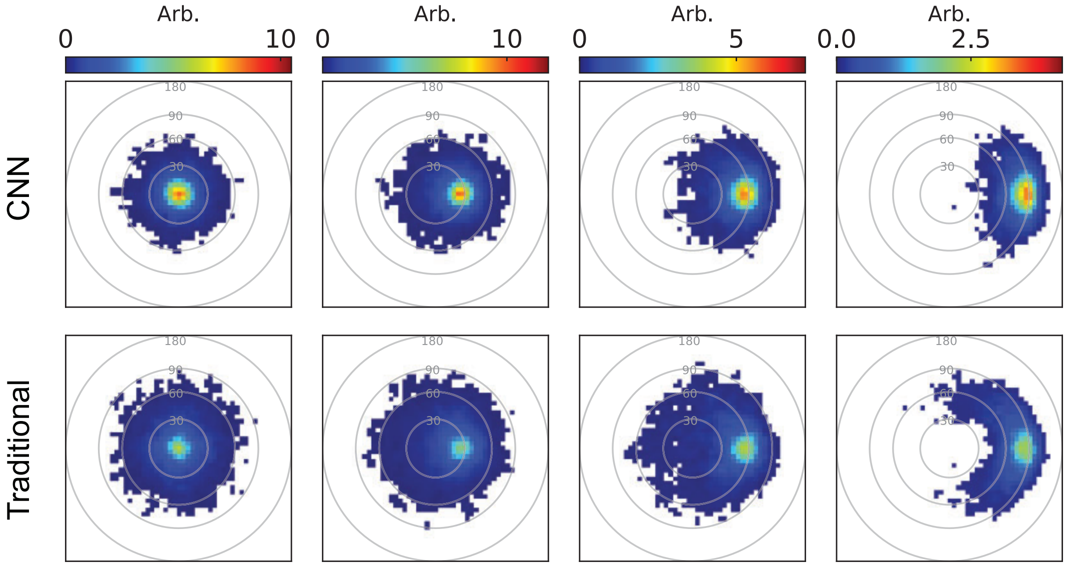

Fig. 10. Reconstructed source images of ^{137}Cs set at zenith angles of 0, 30, 60, and 90° for the experimental data. The top and bottom columns are the CNN and traditional methods, respectively.

Reconstructed source images for ^{137}Cs at several zenith angles are shown in Fig. 10 using an equal-solid-angle projection, such as the Lambert projection. The bright points representing sources were observed, and the more focused images were confirmed in the CNN method.

6. Discussion

Since the scattering position resolution is smaller than the PSA position uncertainty, further improvement in the scattering points cannot be expected to improve the PSF. On the other hand, the SPD has potential for improvement. Although the recoil direction was dramatically improved using CNN, the physical limit of the multiple scattering has not yet been reached. As the angular resolutions of the azimuth and the zenith angles were calculated separately for 50 keV electrons at a 4 mm distance from the scattering points, 29° and 35° were obtained, respectively. This indicates that the zenith angular resolution mainly determines the 3D angular resolution. The Z-direction position resolution has a width corresponding to the TOT. From a rough calculation, a drift speed of 38 mm μs^{-1} and

TOT width of 80 ns corresponding to a 0.5 pC input charge give a 3.0 mm position resolution, which is 3.8 times worse than the 800 μm read-out pitch of the X and Y directions. This is the cause of smearing of the track information for low-energy electrons, as can be seen from the 42.3 keV image in the upper left of Fig. 6. Improvement can be expected by using a short shaping time amplifier. Also, waveforms taken by the flush ADC have Z-direction information [33]. Therefore, utilizing the waveform images in CNNs would be a good approach. Regarding the azimuth direction, it is adequate to make the read-out pitch more acceptable and increase the track information using, for example, a three-projection read-out system [34,35]. We confirmed that a three-projection read-out system with 480 μm pitch achieved 50° of 3D angular resolution from the simulation.

In this research, a network based on the VGG model was used as a demonstration. However, residual networks [19] (ResNet) have mainly been used to realize deeper layers in recent research. In the future, we will incorporate this technology to improve the neural network aspect.

7. Conclusion

We designed two CNNs based on the VGG and U-Net models to predict the scattering position and the electron recoil direction from track images taken by the ETCC. The angular resolution and scattering position resolution utilizing CNNs were obtained as 41° and 2.1 mm for 75 keV simulation data. As we adopted the CNN analysis for experimental calibration data, the resulting PSF was improved by 32% compared with the traditional analysis and achieved 15°. The CNN analysis surpassed the traditional analysis, and the improved ETCC has the potential to reveal the MeV gamma-ray sky.

Acknowledgements

This work was partly supported by Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant-in-Aids for Scientific Research) (grant no. 20K20428).

Funding

Open Access funding: SCOAP^3.

References

[1] S. M. Matz, G. H. Share, M. D. Leising, E. L. Chupp, W. T. Vestrandt, W. R. Purcell, M. S. Strickman, and C. Reppin, Nature **331**, 416 (1988).
[2] R. Diehl, et al., Astron. Astrophys. **411**, L451 (2003).
[3] E. Churazov et al., Astrophys. J. **812**, 62 (2015).
[4] M. Chiaberge, A. Capetti, and A. Celotti, Mon. Not. Roy. Astron. Soc. **324**, L33 (2001).
[5] L. Ballo et al., Astrophys. J. **567**, 50 (2002).
[6] M. S. Briggs et al., Astrophys. J. **524**, 82 (1999).
[7] R. D. Preece, M. S. Briggs, R. S. Mallozzi, G. N. Pendleton, W. S. Paciesas, and D. L. Ban, Astrophys. J. Suppl. Ser. **126**, 19 (2000).
[8] W. R. Purcell, L.-X. Cheng, D. D. Dixon, R. L. Kinzer, J. D. Kurfess, M. Leventhal, M. A. Saunders, J. G. Skibo, D. M. Smith, and J. Tueller, Astrophys. J. **491**, 725 (1997).
[9] T. Siegert, R. Diehl, G. Khachatryan, M. G. H. Krause, F. Guglielmetti, J. Greiner, A. W. Strong, and X. Zhang, Astron. Astrophys. **586**, A84 (2016).
[10] C. A. Kierans et al., Astrophys. J. **895**, 44 (2020).
[11] V. Schönfelder et al., Astron. Astrophys. Supp. Ser. **143**, 145 (2000).
[12] T. Tanimori et al., Sci. Rep. **7**, 41511 (2017).
[13] V. Schönfelder, New Astron. Rev. **48**, 193 (2004).
[14] T. Tanimori et al., Astrophys. J. **810**, 28 (2015).
[15] A. Takada et al., Nucl. Instr. Meth. Phys. Res. A **573**, 195 (2007).

[16] A. Krizhevsky, I. Sutskever, and G.E. Hinton, in *Proc. 25th Int. Conf. Neural Information Processing Syst.*, eds. F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (2012), Vol. 1.

[17] K. Simonyan and A. Zisserman, arXiv:1409.1556 [cs.CV] [Search INSPIRE].

[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, arXiv:1409.4842 [cs.CV] [Search INSPIRE].

[19] K. He, X. Zhang, S. Ren, and J. Sun, in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2016), p. 770.

[20] O. Russakovsky et al., arXiv:1409.0575 [cs.CV] [Search INSPIRE].

[21] T. Kitaguchi, K. Black, T. Enoto, A. Hayato, J. E. Hill, W. B. Iwakiri, P. Kaaret, T. Mizuno, and T. Tamagawa, Nucl. Instr. Meth. Phys. Res. A **942**, 162389 (2019).

[22] M. Kekic et al. [NEXT collaboration], J. High Energy Phys. **2101**, 189 (2021).

[23] T. Tanimori [on behalf of SMILE project], J. Phys.: Conf. Ser. **1468**, 012046 (2020).

[24] T. Mizumoto et al., Nucl. Instr. Meth. Phys. Res. A **800**, 40 (2015).

[25] M. Abadi et al. arXiv:1603.04467 [cs.DC] [Search INSPIRE].

[26] V. Nair and G. E. Hinton, in *Proc. 27th Int. Conf. Int. Conf. Machine Learning*, eds. J. Fürnkranz and T. Joachims. (Omnipress, Madison, WI, 2010), p. 807.

[27] E. Shelhamer, J. Long, and T. Darrell, IEEE Trans. Patt. Anal. Mach. Intel. **39**, 640 (2017).

[28] O. Ronneberger, P. Fischer, and T. Brox, in *Proc. 18th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, eds. N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi (Springer, New York, 2015), Vol. 3, p. 234.

[29] S. Agostinelli et al., Nucl. Instr. Meth. Phys. Res. A **506**, 250 (2003).

[30] S. F. Biagi, Nucl. Instr. Meth. Phys. Res. A **283**, 716 (1989).

[31] A. Takada et al., J. Instrum. **8**, C10023 (2013).

[32] F. Sauli, CERN-77-09 (CERN, Geneva, 2004).

[33] J. B. R. Battat et al., J. Instrum. **11**, P10019 (2016).

[34] F. Sauli, T. Meinschad, L. Musa, and L. Ropelewski, *Conf. Rec. IEEE Symp. Nucl. Sci.*, (IEEE, New York, 2004), Vol. 1, p. 12.

[35] M. Diwan, M. Potekhin, V. Viren, X. Qian, and C. Zhang, J. Phys. Conf. Ser. **762**, 012033 (2016).