Research on X-ray Polarization Demodulation Method Based on Hexagonal Convolutional Neural Network

Authors: Ya-Nan Li, Jia-Huan Zhu, Huai-Zhong Gao, Hong Li, Ji-Rong Cang, Zhi Zeng, Hua Feng, Ming Zeng, Ming Zeng

Date: 2024-07-30T00:00:00+00:00

Abstract

Track reconstruction algorithms are critical for polarization measurements. Convolutional neural networks (CNNs) are a promising alternative to traditional moment-based track reconstruction approaches. However, the hexagonal grid track images obtained using gas pixel detectors (GPDs) for better anisotropy do not match the classical rectangle-based CNN, and converting the track images from hexagonal to square results in a loss of information. We developed a new hexagonal CNN algorithm for track reconstruction and polarization estimation in X-ray polarimeters, which was used to extract the emission angles and absorption points from photoelectron track images and predict the uncertainty of the predicted emission angles. The simulated data from the PolarLight test were used to train and test the hexagonal CNN models. For individual energies, the hexagonal CNN algorithm produced 15%–30% improvements in the modulation factor compared to the moment analysis method for 100% polarized data, and its performance was comparable to that of the rectangle-based CNN algorithm that was recently developed by the Imaging X-ray Polarimetry Explorer team, but at a lower computational and storage cost for preprocessing.

Full Text

Preamble

Research on X-ray Polarization Reconstruction Using Hexagonal Convolutional Neural Networks

Ya-Nan Li[1,2] · Jia-Huan Zhu[3] · Huai-Zhong Gao[1,2] · Hong Li[3] · Ji-Rong Cang[4] · Zhi Zeng[2] · Hua Feng[3] · Ming Zeng[1,2]

[1] Key Laboratory of Particle & Radiation Imaging (Tsinghua University), Ministry of Education, Beijing 100084, China
[2] Department of Engineering Physics, Tsinghua University, Beijing 100084, China
[3] Institute for Advanced Study, Tsinghua University, Beijing 100084, China
[4] StarDetect Co., Ltd., Beijing 100084, China

This work is supported by the National Natural Science Foundation of China (Grant No. 12025301) and the Tsinghua University Initiative Scientific Research Program.

Corresponding author: zengming@tsinghua.edu.cn (Ming Zeng)

Abstract

Track reconstruction algorithms are critical for polarization measurements. Convolutional neural networks (CNNs) represent a promising alternative to traditional moment-based track reconstruction approaches. However, the hexagonal-grid track images produced by gas pixel detectors (GPDs) for improved isotropy are incompatible with conventional rectangle-based CNNs, and converting these images from hexagonal to square grids results in information loss. We developed a novel hexagonal CNN algorithm for track reconstruction and polarization estimation in X-ray polarimeters, which extracts photoelectron emission angles and absorption points from track images while simultaneously predicting the uncertainty of the estimated emission angles. Using simulated data from the PolarLight experiment, we trained and evaluated our hexagonal CNN models. For individual photon energies, the hexagonal CNN algorithm achieved 15%–30% improvements in modulation factor compared to the moment analysis method for 100% polarized data. Its performance was comparable to that of the rectangle-based CNN algorithm recently developed by the Imaging X-ray Polarimetry Explorer (IXPE) team, but with lower computational and storage costs for preprocessing.

Keywords: X-ray polarization, Track reconstruction, Deep learning, Hexagonal convolutional neural network

Introduction

Astronomical X-ray polarimetry is a powerful tool for probing the magnetic fields, geometries, and emission mechanisms of high-energy astrophysical sources [**?**, **?**]. The field originated in the 1960s with attempts to detect soft X-ray polarization from the Crab Nebula, Scorpius X-1, and other objects using Bragg diffraction and Thomson scattering polarimeters [**?**]. However, the limited sensitivity of these early instruments brought astronomical X-ray polarization measurements to a standstill for over 40 years following the OSO-8 satellite experiments.

In the few-keV energy range, the photoelectric effect dominates light-matter interactions. The differential cross-section for photoelectron emission is proportional to $\cos^2 \varphi$, where $\varphi = \varphi_e - \varphi_0$, with φ_e representing the azimuthal angle of the photoelectron and φ_0 the electric vector position angle (EVPA) of the X-ray photon [?]. Consequently, the polarization fraction and EVPA of an X-ray source can be determined by measuring the emission angles of numerous photoelectrons. The advent of micropattern gas detectors has enabled polarimetry based on the photoelectric effect, dramatically improving polarization sensitivity [?]. Modern missions such as the PolarLight CubeSat experiment [?, ?, ?], the Imaging X-ray Polarimetry Explorer (IXPE) [?], the enhanced X-ray Timing and Polarimetry mission (eXTP) [?, ?], and the Cosmic X-ray Polarization Detector (CXPD) [?, ?] all employ gas pixel detectors (GPDs) for polarization detection.

Reconstructing photoelectron emission angles from track images is complicated by Coulomb scattering, transverse diffusion during drift, and electronic noise. The performance of the track reconstruction algorithm directly impacts polarimeter sensitivity. Current approaches fall into two categories: traditional algorithms such as moment analysis [?], adaptive cut methods [?], and graph-based methods [?]; and CNN-based reconstruction methods [?, ?, ?], which leverage powerful image processing capabilities to achieve superior performance. However, major space polarimetry missions utilize hexagonal-pixel application-specific integrated circuits (ASICs) for track readout to enhance isotropy [?, ?, ?, ?, ?, ?], resulting in photoelectron track images with a hexagonal pixel structure (Fig. 1 [Figure 1: see original paper]). Existing CNN-based methods employ classical rectangle-based CNNs, requiring conversion from hexagonal to approximate square-pixel images, which inevitably leads to information loss.

Hexagonal CNNs are deep learning models designed specifically for hexagonal-pixel structures, using hexagonal convolutional kernels instead of rectangular ones to better capture spatial context information. Applying hexagonal CNNs to process hexagonal-pixel photoelectron track images from GPDs is expected to yield improved polarization reconstruction performance. In this study, we propose a novel X-ray polarization reconstruction method based on hexagonal CNNs. The remainder of this paper is organized as follows: Section 2 introduces hexagonal CNNs and uncertainty quantification in deep learning. Section 3 describes the training procedure for hexagonal CNNs in photoelectron track reconstruction. Section 4 presents the prediction and reconstruction results. Finally, Section 5 concludes with a summary and future prospects.

2.1 Hexagonal CNNs

Convolutional neural networks have garnered considerable attention due to their exceptional performance in computer vision and big data applications [?, ?, ?, ?, ?]. As their application domains expand, classical CNNs based on Cartesian architectures have become inadequate for complex problems. Significant advances have been made in network architecture design and convolutional operations, generalizing CNNs for multi-view applications [?], non-Euclidean spaces [?], and other domains.

While images are typically acquired using square sensor arrays, square grids are not optimal for planar sampling. Hexagonal grids offer several advantages over square grids, including six-fold rotational symmetry, a smaller edge-to-area ratio, and equidistant neighboring pixels. These properties make hexagonal grids widely applicable in cosmology, astrophysics, and visual systems. Hexagonal CNNs are deep learning networks built upon hexagonal grids, where hexagonal convolution kernels replace the rectangular kernels of classical CNNs. The differences between these kernels are illustrated in Fig. 2 [Figure 2: see original paper]. Compared to classical CNNs, hexagonal CNNs exhibit better symmetry and unique advantages for aerial imagery and geospatial information [?].

Despite these advantages, hexagonal CNNs suffer from higher computational complexity and training difficulty. Existing research follows two main approaches. The first reuses highly optimized rectangular convolution routines to implement hexagonal convolution, as exemplified by HexagDLy [?] and HexagonNet [?]. The second focuses on native hexagonal CNN architectures that implement hexagonal convolution operations directly, such as HexCNN [?]. While native hexagonal CNNs offer significant advantages in training time and memory efficiency, they cannot yet leverage GPU parallelization for accelerated training and inference [?]. HexagDLy is a Python library that performs convolution and pooling operations on hexagonal pixel data. Fig. 3 [Figure 3: see original paper] illustrates a convolution implementation with a hexagonal kernel size of 1 in HexagDLy [?]. We constructed our hexagonal CNN architecture for track reconstruction using HexagDLy due to its flexibility and user-friendliness, with detailed architecture discussed in Section 3.

2.2 Uncertainty Quantification

Uncertainty estimation in deep learning has become an active research area, enabling neural networks to output not only predictions $\hat{y}$ for input x but also predictive uncertainty $\sigma_{\hat{y}}$, thereby expanding the applicability of deep learning. Uncertainty estimation is particularly valuable for X-ray polarization reconstruction, as demonstrated in [?, ?].

Two primary uncertainty types exist in deep learning: aleatoric uncertainty (data uncertainty) and epistemic uncertainty (model uncertainty) [?]. Aleatoric uncertainty captures inherent data noise and class overlap that cannot be reduced by collecting more data. Epistemic uncertainty arises from limited knowledge of the data distribution or inadequate model structure and can theoretically be reduced through more complex models, expanded datasets, or regularization techniques [**?**].

Aleatoric uncertainty can be modeled by augmenting the loss function. Assuming data noise follows a Gaussian distribution ($\varepsilon \sim N(0, \sigma^2)$), the model's predicted output distribution for a given input x is $N(\hat{y}, \sigma^2)$. The loss function for predicting aleatoric uncertainty σ_a is obtained by minimizing the negative log-likelihood (NLL) loss across all training data (Eq. 1):

$$L(y_i|x_i) = \frac{\|y_i - \hat{y}(x_i)\|^2}{2\sigma_{\hat{a}}^2(x_i)} + \log(\sigma_{\hat{a}}(x_i))$$

Epistemic uncertainty is significantly more challenging to quantify, though numerous methods have been proposed, including Bayesian neural networks (BNNs), deep ensembles, and evidential deep regression (EDR). BNNs model epistemic uncertainty by placing prior distributions on network weight parameters ω and deriving posterior distributions $P(\omega|D)$ from dataset D. However, BNNs are difficult to apply in practice because the posterior distribution $P(\omega|D)$ is typically intractable. Approximation algorithms such as variational inference and Monte Carlo dropout have been proposed to estimate epistemic uncertainty.

Deep ensembles represent another powerful approach for modeling epistemic uncertainty and have been widely adopted. A deep ensemble uses multiple base models to generate M predictions $\{\hat{y}_j\}$ for the same input x, where M is the number of models. The variance across these predictions serves as the epistemic uncertainty estimate. Deep ensembles are straightforward to implement and can achieve uncertainty estimation performance comparable to or better than BNN approximation algorithms.

Evidential deep regression directly learns higher-order distributions of network outputs, using a deterministic network to learn both aleatoric and epistemic uncertainties by placing evidential priors over the original aleatoric uncertainty loss function. While EDR has shown promising results, recent studies have identified theoretical shortcomings in its mathematical foundations [**?**].

Considering these factors, we chose to estimate aleatoric uncertainty through loss function augmentation and epistemic uncertainty using deep ensembles, which offer greater stability and easier implementation.

---

3 Hexagonal CNN Model Training for Photoelectron Track Reconstruction

X-ray polarization reconstruction algorithms typically involve two steps. First, track features are extracted from individual photoelectron track images (track reconstruction), including photoelectron emission angles (φ), absorption points (x, y), and photoelectron energies (E). Second, the polarization parameters (polarization fraction and EVPA) of the X-ray source are estimated from the emission angles of numerous photoelectrons. Among these, extracting track features from blurred photoelectron track images is the key to successful polarization reconstruction. This section details our implementation of the hexagonal CNN-based photoelectron track feature reconstruction algorithm.

3.1 Dataset

We employed supervised learning for photoelectron track reconstruction. Since the true emission angle φ in experimental data follows a $\cos^2\varphi$ distribution and is therefore unknown, the training dataset had to be generated through simulation.

PolarLight is a small X-ray GPD onboard a CubeSat that conducts on-orbit scientific observations. It uses an ASIC designed by the INFN-Pisa group with a pixel matrix of 352×300 (105k pixels) and hexagonal pixels with a 50 m pitch for track readout [?]. The PolarLight experiment utilized Monte Carlo Geant4/Garfield simulations validated against experimental data. We used this simulation framework to generate our photoelectron track dataset.

The photoelectron track features include emission angles, absorption points, and photoelectron energy. Since photoelectron energy reconstruction is relatively straightforward and can be accomplished effectively with non-CNN algorithms, this study focused exclusively on reconstructing (φ, x, y).

The CNN training dataset must be uniformly distributed to prevent overfitting, low prediction accuracy, or biased results. Therefore, we configured the incident X-ray parameters in our simulation as follows: (1) To ensure uniform emission angle distribution, we set the incident X-ray polarization to zero (unpolarized). (2) To ensure the hexagonal CNN model performed well across the entire detector plane, we uniformly distributed the incident X-ray coordinates (x, y). (3) Since PolarLight's effective energy range is 2–8 keV, we uniformly distributed incident X-ray energies from 2–9 keV to ensure adequate training at the energy interval boundaries. Low-energy photoelectron track images are excessively noisy, making feature extraction difficult, so we did not extend the dataset to lower energies.

We simulated 870,050 photoelectron tracks with uniform distributions of emission angles, absorption points, and energies, splitting them into training (90%), validation (5%), and test (5%) sets.

Photoelectron tracks originate not only from photon interactions within the

GPD gas but also from interactions with detector components outside the gas volume (e.g., the beryllium window and gas electron multiplier (GEM)), which create low-energy tails in the energy histogram [**?**, **?**]. These tracks are difficult to reconstruct and were excluded from our dataset. Additionally, high-energy photoelectrons may traverse the detector without depositing their full energy; these tracks were also removed.

Since HexagDLy implements hexagonal convolution based on rectangular convolution routines, image preprocessing was required to convert hexagonal-grid tracks into square-grid images. Fig. 4 [Figure 4: see original paper] shows an example of this preprocessing. To fully exploit the six-fold rotational symmetry of hexagonal-grid images, each photoelectron track image was converted into three input images: unrotated, rotated by +60°, and rotated by -60°.

Considering the photoelectron track length and the readout ASIC pixel size, we set the photoelectron track image size to 64×64.

3.2 Loss Function

Track reconstruction is a multi-task problem requiring separate loss functions for emission angles and absorption points.

Photoelectron emission angles are periodic and their distribution follows a Von Mises (VM) distribution—a continuous probability distribution on $[0, 2\pi]$ that serves as the circular analog of the normal distribution. For predicting epistemic uncertainty, the Von Mises NLL is more appropriate than the Gaussian NLL described in Section 2.2. The loss function for emission angles based on the Von Mises distribution for a single hexagonal CNN model is given by Eq. 2, with detailed derivation in [**?**]:

$$L_{\varphi}(v|x) = -\hat{\kappa}_a(\hat{v}_2 \cdot v_2) + \log I_0(\hat{\kappa}_a) \quad (2)$$

where I_0 is the modified Bessel function of the first kind (order 0), $v_2 = (\cos 2\varphi, \sin 2\varphi)$ accounts for the 2φ dependence of X-ray polarization, and non-negative $\hat{\kappa}_a$ is the predicted aleatoric VM uncertainty parameter. The circular variance of the VM distribution can be derived from $\hat{\kappa}_a$ using Eq. 3, where I_1 is the modified Bessel function of the first kind (order 1):

$$\frac{I_1(\hat{\kappa}_a)}{I_0(\hat{\kappa}_a)}$$

Assuming the epistemic VM uncertainty κ_e also follows a Von Mises distribution $VM(0, \kappa_e)$, $\hat{\kappa}_e$ can be estimated from the emission angles of M hexagonal CNN models in a deep ensemble [**?**]:

$$\bar{R}_2 = \left(\frac{1}{M}\sum_{j=1}^{M}\cos 2\hat{\varphi}_j\right)^2 + \left(\frac{1}{M}\sum_{j=1}^{M}\sin 2\hat{\varphi}_j\right)^2 \quad (4)$$

$$\frac{I_1(\hat{\kappa}_e)}{I_0(\hat{\kappa}_e)} = \bar{R} \quad (5)$$

The total uncertainty variance σ^2 of emission angle φ is obtained by summing the aleatoric uncertainty variance σ_a^2 and epistemic uncertainty variance σ_e^2. The total error σ_i of emission angle prediction φ_i for track x_i is then given by Eq. 6, where a factor of 1/2 transforms the errors from $2\varphi_i$ to φ_i:

$$\sigma_i = \sqrt{\frac{1}{2}\left(1 - \frac{I_1(\hat{\kappa}_{ij})}{I_0(\hat{\kappa}_{ij})}\right) + \frac{1}{2}\left(1 - \frac{I_1(\hat{\kappa}_i)}{I_0(\hat{\kappa}_i)}\right)} \quad (6)$$

The loss function for absorption points is the standard L2 loss (Eq. 7). While absorption point uncertainty is not the focus of this study, it can be obtained using Eq. 1 combined with a deep ensemble if needed:

$$L_{xy}(x_0, y_0|x) = \|(x_0, y_0) - (\hat{x}(x), \hat{y}(x))\|^2 \quad (7)$$

The total loss function for a single hexagonal CNN model is:

$$L = L_\varphi + \alpha\|v_1 - \hat{v}_1\|^2 + \beta L_{xy} + \sigma\|\omega\|^2 \quad (8)$$

where $\|v_1 - \hat{v}_1\|^2$ enables prediction of emission angles across the full 2π range, and the final term $\sigma\|\omega\|^2$ prevents overfitting. Three hyperparameters (α, β, σ) weight the individual loss components. Each hexagonal CNN model predicts a five-dimensional vector $(\cos\varphi, \sin\varphi, \kappa, x, y)$ for a given track image input x.

3.3 Hexagonal CNN Architecture

Reconstructing emission angles and absorption points is a complex task that cannot be satisfied by simple architectures with only three or four convolutional layers. Considering the characteristics of photoelectron tracks, we built our hexagonal CNN based on the ResNet-18 architecture [**?**].

A residual block forms the basic unit of a residual network. Our hexagonal residual block (Fig. 5 [Figure 5: see original paper]) was constructed using HexagDLy's hexagonal convolution operations, comprising a hexagonal convolution layer with kernel size 1, batch normalization, ReLU activation, another hexagonal convolution layer with kernel size 1, and batch normalization, followed by a skip connection that bypasses these layers and feeds directly into the final ReLU activation. These hexagonal residual blocks are repeated to form the complete hexagonal CNN architecture for track reconstruction.

Table 1 details the architecture. The conv1 layer uses a hexagonal convolution (kernel size = 1) and hexagonal max pooling (kernel size = 1, stride = 2) to extract initial track features. The conv2–conv5 layers, composed of hexagonal residual blocks, extract deeper features. Finally, the feature maps from conv5 are converted into the five-dimensional output vector $(\cos\varphi, \sin\varphi, \kappa, x, y)$ through an average pooling layer and a fully connected layer.

3.4 Training

Standardization was applied to training data to prevent vanishing/exploding gradients and accelerate convergence:

$$x_{\text{norm}} = \frac{x-\mu}{\sigma}$$

where μ and σ are the pixel mean and standard deviation computed across the entire training set.

The hexagonal CNN model was optimized using stochastic gradient descent with momentum (SGD), a standard deep learning optimization algorithm. The learning rate was stepwise-decayed from an initial value of 0.005. Model parameters were randomly initialized before training, with five differently initialized models generated for deep ensembles. Considering memory constraints, batch sizes of 512 and 1024 were used. Training proceeded for 150 epochs with loss function hyperparameters set to $\alpha = 0.3$, $\beta = 0.2$, and $\sigma = 5 \times 10^{-5}$.

---

4 Results

This section reports the performance of our hexagonal CNN method for track and polarization reconstruction using simulated PolarLight track images.

4.1 Emission Angle Reconstruction and Uncertainty Estimation

Accurate photoelectron emission angle reconstruction forms the foundation of X-ray polarization measurement. Improved reconstruction accuracy directly enhances polarization reconstruction performance.

Our hexagonal CNN method predicts emission angle φ using $(\cos\varphi, \sin\varphi)$. As described in Section 3, image rotation augmentation converted each track image into three input images. Consequently, each hexagonal CNN model produced three prediction sets per track image. With $M = 5$ models in our deep ensemble, each track yielded $3M$ predictions. The final emission angle for a single track was calculated using Eq. 10:

$$\bar{\varphi}_i = \arctan 2\left(\sum_{j=1}^{3M} \sin\varphi_{ij}, \sum_{j=1}^{3M} \cos\varphi_{ij}\right) \quad (10)$$

Fig. 6 [Figure 6: see original paper] compares emission angle reconstruction results from moment analysis and hexagonal CNN methods for 3 keV and 9 keV photoelectrons. Due to difficulty distinguishing the track's beginning from its end—particularly at lower energies—a 180° ambiguity appears as two sub-bright lines parallel to the central bright line. Notably, this ambiguity does not affect polarization reconstruction, where EVPA ranges from $-\pi/2$ to $+\pi/2$. The hexagonal CNN method demonstrates superior ability to distinguish track endpoints and achieves higher reconstruction accuracy compared to moment analysis.

Root mean square error (RMSE) was calculated using Eq. 11 to evaluate reconstruction accuracy:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\varphi}_i - \varphi_i)^2} \quad (11)$$

where $\hat{\varphi}-\varphi$ is folded into the range $[-\pi/2, \pi/2]$ since polarization reconstruction depends only on 2φ.

Fig. 7 [Figure 7: see original paper] shows RMSE as a function of incident X-ray energy for both methods on unpolarized PolarLight data. For low-energy photoelectrons with short, noisy tracks, the hexagonal CNN method did not significantly outperform moment analysis. As energy increased, tracks became longer with clearer initial segments, and both methods achieved high accuracy, but the hexagonal CNN method showed significant superiority for these complex tracks.

Another crucial prediction is the emission angle error σ_i, calculable from the hexagonal CNN's κ output (Eq. 6). This predicted uncertainty helps filter events with poor reconstruction, improving polarization estimation effectiveness. While quantitative validation of prediction accuracy is challenging, we assessed reasonableness by comparing predicted errors with factors influencing reconstruction difficulty.

Emission angle reconstruction difficulty correlates with transverse electron diffusion during drift, track shortening from 3D-to-2D projection, and photoelectron energy. We established a coordinate system for PolarLight's effective sensitive volume ($15\,\text{mm} \times 15\,\text{mm} \times 10\,\text{mm}$), defining the x-y plane as the readout plane, z-axis as the electric field direction, and the volume center as $(0, 0, 0)$.

Fig. 8 [Figure 8: see original paper] shows predicted error distributions as functions of these factors. Fig. 8(a) displays error versus drift distance d (distance from photon absorption point to GEM) for 7 keV photoelectrons. Transverse diffusion standard deviation follows $\sigma_{\text{drift}} = \sigma_f \sqrt{d}$, where σ_f is the GPD gas diffusion coefficient. Predicted errors increase with drift distance as diffusion worsens. Fig. 8(b) shows error versus photoelectron scattering angle θ (angle between incident X-ray and photoelectron directions) for 3 keV photoelectrons. Tracks more parallel to the readout plane project longer, yielding smaller predicted errors. Fig. 8(c) shows error versus photoelectron energy; higher energies produce longer tracks with clearer initial segments, reducing predicted errors.

These results demonstrate that the hexagonal CNN's predicted errors are reasonable, accurately reflecting track blurring and reconstruction difficulty. Fig. 9 [Figure 9: see original paper] further shows the relationship between actual reconstruction error $|\hat{\varphi} - \varphi|$ and predicted error, confirming that larger actual errors correspond to greater predicted uncertainty.

4.2 Absorption Point Reconstruction

Absorption point reconstruction is crucial for improving polarimeter spatial resolution. Accuracy is evaluated using half-power diameter (HPD), a standard metric in X-ray imaging defined as the diameter of a circle centered on the true absorption point that contains 50% of reconstructed points. Larger HPD indicates poorer reconstruction. Fig. 10 [Figure 10: see original paper] compares HPD values for moment analysis and hexagonal CNN methods across energies. The hexagonal CNN method outperforms moment analysis, particularly for complex high-energy tracks.

4.3 Polarization Estimation

Polarization reconstruction performance directly affects polarimeter sensitivity. We analyzed our hexagonal CNN method's performance and compared it with moment analysis and the rectangle-based CNN method developed by the IXPE team.

Polarized and unpolarized simulation tracks were generated using PolarLight simulation algorithms, with incomplete energy deposition and non-gas interactions removed, consistent with training data. Fig. 11 [Figure 11: see original paper] shows binned modulation curves from predicted emission angles for unpolarized and 100% polarized data at 3 keV and 7 keV. The hexagonal CNN method's residual systematic modulation is as flat as moment analysis, indicating no additional systematic errors, while recovering significantly more modulation from polarized data.

An unbinned polarization estimation algorithm based on Stokes parameters was used to estimate polarization fraction and EVPA from predicted track angles. Fig. 12 [Figure 12: see original paper] compares modulation responses for moment analysis, the IXPE team's rectangle-based CNN method (S-CNN), and our hexagonal CNN method. Our method outperformed moment analysis by 15%–30% in modulation factor across individual energies, achieving performance comparable to the IXPE team's rectangle-based CNN.

This comparable performance may stem from the rectangle-based CNN's double-channel input images compensating for hexagonal-to-square conversion losses, or because existing neural network methods have approached the fundamental limit imposed by photoelectron track blurring, which is difficult to surpass through architectural improvements alone. Nevertheless, our hexagonal CNN method represents a more direct approach to hexagonal-grid track processing, halving input data volume and reducing preprocessing computational and storage costs. While the IXPE method converts each image into three double-channel inputs, ours uses three single-channel inputs. However, since current hexagonal convolution implementations rely on rectangular convolution, memory consumption remains high and could be improved through native hexagonal CNN architectures.

5 Conclusion

We developed a track reconstruction and polarization estimation algorithm based on hexagonal CNNs to match the hexagonal-grid tracks in GPDs for X-ray polarization measurements. Our method predicts photoelectron emission angles, absorption points, and emission angle uncertainties from X-ray track images. The predicted absorption points enable image reconstruction, while predicted emission angles and uncertainties facilitate X-ray source polarization estimation.

Testing with simulated PolarLight data demonstrated that the hexagonal CNN method achieves better absorption point reconstruction (lower HPD) than moment analysis and provides 15%–30% improvements in modulation factor. Performance is comparable to the IXPE team's rectangle-based CNN method but with lower preprocessing computational and storage costs. Our hexagonal CNN method establishes a solid foundation for developing polarization reconstruction algorithms for future eXTP missions.

References

[1] T. Kallman, Astrophysical motivation for X-ray polarimetry. Adv. Space Res. 34, 2673–2677 (2004). https://doi.org/10.1016/j.asr.2003.03.059

[2] P. Meszaros, R. Novick, A. Szentgyorgyi et al., Astrophysical implications and observational prospects of X-ray polarimetry. Astrophys. J. 324, 1056 (1988). https://doi.org/10.1086/165962

[3] R. Giacconi, H. Gursky, F.R. Paolini et al., Evidence for x rays from sources outside the solar system. Phys. Rev. Lett. 9, 439–443 (1962). https://doi.org/10.1103/physrevlett.9.439

[4] M. Gavrila, Relativistic K-shell photoeffect. Phys. Rev. 113, 514–526 (1959). https://doi.org/10.1103/physrev.113.514

[5] E. Costa, P. Soffitta, R. Bellazzini et al., An efficient photoelectric X-ray polarimeter for the study of black holes and neutron stars. Nature 411, 662–665 (2001). https://doi.org/10.1038/35079508
[6] H. Feng, H. Li, X.Y. Long et al., Re-detection and a possible time variation of soft X-ray polarization from the Crab. Nat. Astron. 4, 511–516 (2020). https://doi.org/10.1038/s41550-020-1088-1
[7] X.Y. Long, H. Feng, H. Li et al., A significant detection of X-ray Polarization in Sco X-1 with PolarLight and constraints on the corona geometry.: 2112.02837. (2021).
[8] H. Feng, H. Li, L.D. Kong et al., X-ray polarimetry of the accreting pulsar 1A 0535+262 in the supercritical state with PolarLight. Astrophys. J. 950, 76 (2023). https://doi.org/10.3847/1538-4357/acd0af
[9] M.C. Weisskopf, B. Ramsey, S.L. O'Dell et al., The imaging X-ray polarimetry explorer (IXPE). Results Phys. 6, 1179–1180 (2016). https://doi.org/10.1016/j.rinp.2016.10.021
[10] S.N. Zhang, M. Feroci, A. Santangelo et al., eXTP: enhanced X-ray Timing and Polarimetry Mission.: 1607.08823. (2016).
[11] S.N. Zhang, A. Santangelo, M. Feroci et al., The enhanced X-ray Timing and Polarimetry mission—eXTP. Sci. China Phys. Mech. Astron. 62, 29502 (2018). https://doi.org/10.1007/s11433-018-9309-2
[12] H. Wang, D. Wang, R. Chen et al., Electronics system for the cosmic X-ray polarization detector. Nucl. Sci. Tech. 34, 64 (2023). https://doi.org/10.1007/s41365-023-01221-3
[13] Z.W. Fan, H.B. Liu, H.B. Feng et al., Front-end electronics of CXPD for measuring transient X-ray sources. IEEE Trans. Nucl. Sci. 70, 1507–1513 (2023). https://doi.org/10.1109/TNS.2023.3269091
[14] R. Bellazzini, F. Angelini, L. Baldini et al., Novel gaseous X-ray polarimeter: Data analysis and simulation. Proc. SPIE 4843, Polarimetry in Astronomy (2003). doi: 10.1117/12.459381
[15] T. Kitaguchi, K. Black, T. Enoto et al., An optimized photoelectron track reconstruction method for photoelectric X-ray polarimeters. Nucl. Instrum. Meth. Phys. Res. Sect. A Accel. Spectrometers Detect. Assoc. Equip. 880, 188–193 (2018). https://doi.org/10.1016/j.nima.2017.10.070
[16] T.L. Li, M. Zeng, H. Feng et al., Electron track reconstruction and improved modulation for photoelectric X-ray polarimetry. Nucl. Instrum. Meth. Phys. Res. Sect. A Accel. Spectrometers Detect. Assoc. Equip. 858, 62–68 (2017). https://doi.org/10.1016/j.nima.2017.03.050
[17] T. Kitaguchi, K. Black, T. Enoto et al., A convolutional neural network approach for reconstructing polarization information of photoelectric X-ray polarimeters. Nucl. Instrum. Meth. Phys. Res. Sect. A Accel. Spectrometers Detect. Assoc. Equip. 942, 162389 (2019). https://doi.org/10.1016/j.nima.2019.162389
[18] A.L. Peirson, R.W. Romani, H.L. Marshall et al., Deep ensemble analysis for Imaging X-ray Polarimetry. Nucl. Instrum. Meth. Phys. Res. Sect. A Accel. Spectrometers Detect. Assoc. Equip. 986, 164740 (2021). https://doi.org/10.1016/j.nima.2020.164740

[19] A.L. Peirson, Neural network analysis of X-ray polarimeter data. Handbook of X-ray and Gamma-ray Astrophysics. Singapore: Springer, 2024: 5781-5828. https://doi.org/10.1007/978-981-19-6960-7_{144}
[20] C. Szegedy, W. Liu, Y.Q. Jia et al., Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA. IEEE, 1–9 (2015).
[21] Y.H. Chen, W. Li, C. Sakaridis et al., Domain adaptive faster R-CNN for object detection in the wild. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 3339–3348 (2018).
[22] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA. IEEE, 3431–3440 (2015).
[23] Y.D. Zeng, J. Wang, R. Zhao et al., Decomposition of fissile isotope antineutrino spectra using convolutional neural network. Nucl. Sci. Tech. 34, 79 (2023). https://doi.org/10.1007/s41365-023-01229-9
[24] F.D. Qin, H.Y. Luo, Z.Z. He et al., Counting of alpha particle tracks on imaging plate based on a convolutional neural network. Nucl. Sci. Tech. 34, 37 (2023). https://doi.org/10.1007/s41365-023-01190-7
[25] Y.F. Feng, Z.Z. Zhang, X.B. Zhao et al., GVCNN: group-view convolutional neural networks for 3D shape recognition. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 264–272 (2018).
[26] M.M. Bronstein, J. Bruna, Y. LeCun et al., Geometric deep learning: going beyond euclidean data. IEEE Signal Process. Mag. 34, 18–42 (2017). https://doi.org/10.1109/MSP.2017.2693418
[27] J.T. Ke, H. Yang, H.Y. Zheng et al., Hexagon-based convolutional neural network for supply-demand forecasting of ride-sourcing services. IEEE Trans. Intell. Transp. Syst. 20, 4160–4173 (2019). https://doi.org/10.1109/TITS.2018.2882861
[28] C. Steppa, T.L. Holch, HexagDLy—Processing hexagonally sampled data with CNNs in PyTorch. SoftwareX 9, 193–198 (2019). https://doi.org/10.1016/j.softx.2019.02.010
[29] J.R. Luo, W.P. Zhang, J.M. Su et al., Hexagonal convolutional neural networks for hexagonal grids. IEEE Access 7, 142738–142749 (2019). https://doi.org/10.1109/ACCESS.2019.2944766
[30] Y.X. Zhao, Q.H. Ke, F. Korn et al., HexCNN: a framework for native hexagonal convolutional neural networks. 2020 IEEE International Conference on Data Mining (ICDM). Sorrento, Italy. IEEE, 1424–1429 (2020).
[31] M. Abdar, F. Pourpanah, S. Hussain et al., A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Inf. Fusion 76, 243–297 (2021). https://doi.org/10.1016/j.inffus.2021.05.008
[32] K. Zou, Z.H. Chen, X.D. Yuan et al., A review of uncertainty estimation and its application in medical imaging. Meta-Radiology 1, 100003 (2023). https://doi.org/10.1016/j.metrad.2023.100003
[33] N. Meinert, J. Gawlikowski, A. Lavin, The unreasonable effectiveness of deep evidential regression. Proc. AAAI Conf. Artif. Intell. 37, 9134–9142 (2023). https://doi.org/10.1609/aaai.v37i8.26096

[34] R. Bellazzini, G. Spandre, M. Minuti et al., Direct reading of charge multipliers with a self-triggering CMOS analog chip with 105k pixels at 50 m pitch. Nucl. Instrum. Meth. Phys. Res. Sect. A Accel. Spectrometers Detect. Assoc. Equip. 566, 552–562 (2006). https://doi.org/10.1016/j.nima.2006.07.036
[35] K.M. He, X.Y. Zhang, S.Q. Ren et al., Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 770–778 (2016).

Note: Figure translations are in progress. See original paper for figures.