

Segmentation and Celestial Mapping of Unobservable Regions in Nighttime All-sky Images for the Mephisto Observations

Jian Cui¹, Guo-Wang Du^{1,2*}, Xin-Zhong Er^{3**}, Chu-Xiang Li³, Jun-Fan Hou³, Yu-Xin Xin^{4,5},
Xiang-kun Liu^{1,2}, Xiao-Wei Liu^{1,2}

¹ South-Western Institute for Astronomy Research, Yunnan University, Kunming, Yunnan 650504, China

² Key Laboratory of Survey Science of Yunnan Province, Yunnan University, Kunming, Yunnan 650500,
People's Republic of China

³ Tianjin Astrophysics Center, Tianjin Normal University, Tianjin, 300387, China

⁴ Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011, China;

⁵ Key Laboratory for the Structure and Evolution of Celestial Objects, Yunnan Observatories, Kunming
650011, China

Received 20XX Month Day; accepted 20XX Month Day

Abstract Accurate identification of unobservable regions in nighttime is essential for autonomous scheduling and data quality control in observations. Traditional methods—such as infrared sensing or photometric extinction—provide only coarse, non-spatial estimates of sky clarity, making them insufficient for real-time decision-making. This not only wastes observing time but also introduces contamination when telescopes are directed toward cloud-covered or moonlight-affected regions. To address these limitations, we propose a deep learning-based segmentation framework that provides pixel-level masks of unobservable areas using all-sky images. Supported by a manually annotated dataset of nighttime images, our method enables precise detection of cloud- and moonlight-affected regions. The segmentation results are further mapped to celestial coordinates through Zenithal Equal-Area projection, allowing seamless integration with observation control systems (OCS) for real-time cloud-aware scheduling. While developed for the Mephisto telescope, the framework is generalizable and applicable to other wide-field robotic observatories equipped with all-sky monitoring.

Key words: all-sky camera — deep learning — unobservable regions — segmentation — unattended sky survey

1 INTRODUCTION

The rapid development of time-domain astronomy has been driven by the advent of wide-field, high-cadence sky surveys. A large number of transients and variables have been detected and studied in details. In response to these scientific demands, a new generation of large-scale survey telescopes has emerged. Notable examples include the Zwicky Transient Facility (ZTF), which uses a panoramic camera for high cadence wide-area optical monitoring (Dekany et al. 2020); the Wide Field Survey Telescope (WFST, “Mozi”) at Saishiteng Mountain Observatory in Qinghai, China (Wang et al. 2023); and the Multichannel Photometric Survey Telescope (Mephisto) at Yunnan University’s Lijiang Observatory, which allows simultaneous multiband imaging, i.e., $u/g/i$ or $v/r/z$ (Yuan et al. 2020). The operational systems are designed for robotic operation, rapid follow-up observations automatically, and continuous data acquisition.

Space-borne missions, such as the Einstein Probe (EP) and the multi-band astronomical Variable Objects Monitor (SVOM), are powerful tools for detecting high-energy transients and variable sources (Cordier et al. 2018; Yuan et al. 2025). However, only with the support of rapid multi-band observations can we study the mechanisms of these events (Zhang et al. 2025; Liu et al. 2025; Yin et al. 2025; Du et al. 2025). While satellite observations are immune to terrestrial weather, the immediate follow-up observations by ground-based telescopes rely on weather conditions. Consequently, real-time monitoring of atmospheric conditions is essential to maximize observing efficiency and minimize missed scientific opportunities, especially for robotic infrastructure. Environmental awareness—particularly the detection of clouds and scattered moonlight—is crucial for maintaining data quality and optimizing telescope scheduling.

Atmospheric effects, primarily cloud cover and scattered moonlight, cause regions of the sky that are unobservable and can severely degrade photometric measurements in automated surveys. These conditions obscure sources, distort backgrounds and point spread functions, and undermine both data quality and calibration. Early identification and avoidance of such regions are therefore vital for preserving survey efficiency and scientific integrity.

Different techniques have been developed and deployed to assess nighttime sky observability. Infrared cloud sensors (IRCS) utilize thermal contrasts to infer cloudiness and provide rapid, observatory-wide assessments (Takato et al. 2003; Lewis et al. 2010; Redman et al. 2018; Xin et al. 2020). Ground based sky imagers (SIs), originally developed in atmospheric science, enable real-time cloud mapping with visible and infrared cameras, aided by robust image-processing pipelines (Cazorla et al. 2008; Ghonima et al. 2012; Shields et al. 2013). By comparing observed stellar photometry with catalog, one can estimate the atmospheric transmission spatially, and schedule the observation accordingly (Ofek et al. 2012; Chambers et al. 2016; Tonry et al. 2018). Classical image-processing techniques—including thresholding, histogram equalization, and morphological filtering—offer computational efficiency and have been broadly adopted for cloud segmentation tasks (Souza-Echer et al. 2006; Calbó & Sabburg 2008; Kazantzidis et al. 2012; Taravat et al. 2015). While these approaches have enabled significant progress and remain valuable baselines, several key challenges persist, including precise segmentation at the pixel level, robustness under varying illumination or complex backgrounds, and timely response for modern automated observatory con-

* email: dugking@ynu.edu.cn

** email: phioen@163.com

trol. Addressing these limitations motivates further exploration of advanced methods with operational efficiency in challenging nighttime conditions.

These challenges underscore the need for data-driven frameworks capable of robust, pixel-level segmentation of unobservable regions in nighttime all-sky images. Recent advances in machine learning—particularly convolutional neural networks (CNNs)—have demonstrated great potential in astronomical applications such as transient detection and image analysis (Duev et al. 2019; Muthukrishna et al. 2019; Hausen & Robertson 2020; Forster 2021; Fu et al. 2024; Jia et al. 2023; Sun et al. 2023). Semantic segmentation models like U-Net have proven effective in delivering pixel-level discrimination in complex nighttime sky imagery, consistently outperforming threshold-based or empirical methods (Li et al. 2012; Liu et al. 2015; Cheng & Lin 2017; Xie et al. 2020; Fabel et al. 2022). Nevertheless, several challenges remain: complex cloud morphology, weak illumination, and interference from stars or artificial light complicate both training and annotation, while hardware constraints limit real-time deployment in remote observatories. Furthermore, segmentation outputs in image coordinates are not directly compatible with observatory control systems (OCS), hindering their seamless integration into observation planning.

The Mephisto project, initiated by Yunnan University in 2018, focuses on high-cadence, wide-field, and autonomous time-domain surveys (Yuan et al. 2020). However, in its current phase of unattended operation, Mephisto’s dynamic scheduling and robust target execution are frequently hindered by nighttime atmospheric effects—especially cloud cover and scattered moonlight—that degrade photometric quality but often escape detection by conventional monitors.

In this work, we propose a deep learning-based framework tailored for unattended, wide-field telescopes to achieve robust, real-time segmentation of unobservable sky regions in all-sky images. Our approach leverages recent advances in semantic segmentation and incorporates a coordinate transformation pipeline, enabling seamless integration with observatory control systems. Although developed and validated on Mephisto, the framework is broadly applicable to other survey facilities equipped with all-sky cameras, providing a scalable and practical foundation for visibility-aware scheduling and fully autonomous operations in time-domain astronomy. Section 2 describes the all-sky camera setup, data preprocessing, and annotation protocol at Lijiang Observatory. Section 3 details our segmentation framework, including the improved U-Net model and its training. Section 4 outlines the procedure for mapping segmentation results to celestial coordinates for operational scheduling. Finally, Section 5 summarizes our main result and discusses future plans.

2 ALL-SKY CAMERA SYSTEM AND NIGHTTIME IMAGE DATASET

The all-sky camera facility is installed at Lijiang Observatory (IAU code: 044), Yunnan Astronomical Observatories, Chinese Academy of Sciences (CAS), situated at $100^{\circ}01'48''$ east longitude, $26^{\circ}41'42''$ north latitude, and an altitude of approximately 3,200 meters (Wang et al. 2019). The site benefits from minimal air pollution and low dust content, resulting in a clean and transparent atmosphere that provides excellent conditions for astronomical observations (Tan et al. 2002). It is equipped with a high-sensitivity, low-noise CMOS sensor paired with a fisheye lens, providing a 360° field of view for comprehensive sky coverage. It is capable of capturing high-resolution images in real time—even under extremely low-light

nighttime conditions—and supports high-frequency image acquisition as well as real-time data transmission. These features make the facility particularly effective for long-term sky survey operations. While the system enables continuous, all-weather monitoring of the entire sky, regular maintenance and human supervision are still necessary to ensure stable performance and reliable data quality (Xin et al. 2020).

To construct a standardized nighttime sky dataset, 2,000 representative images are systematically selected from the all-sky camera data collected between January and September 2024, specifically during nighttime hours (20:00 to 04:00). This dataset covers a comprehensive array of nighttime weather conditions observable by the all-sky camera, including clear skies, thin or scattered clouds, partially cloudy conditions, and fully overcast skies. By encompassing this diversity of atmospheric phenomena captured during nighttime observations, the dataset ensures both temporal variability and meteorological representativeness. The raw images, as illustrated in Fig. 1(a), initially exhibit variations in dimensions (4144×2822 pixels) in 24-bit RGB color format. To ensure consistency and maintain high data quality, a dedicated preprocessing workflow is applied.

1. **Image Cropping and Resizing:** Each image is cropped to the largest inscribed circle containing cloud features and then resized to 2568×2568 pixels, thereby preserving essential sky and cloud structures while standardizing the input size. Fig. 1(b) shows the cropped and resized circular region.
2. **Annotation:** The cropped images are manually annotated using image editing software, where obstructed or irrelevant regions are masked to generate precise ground-truth labels. During annotation, static regions such as the ground, observatory dome, and DIMM tower are masked as “unobservable regions” to maintain a consistent data format. Since these structures occupy fixed positions and do not obstruct actual telescope observations, they are excluded from subsequent analysis by applying a static mask template. As a result, only dynamic atmospheric factors—such as clouds and moonlight—are considered in visibility assessment and observation scheduling. Fig. 1(c) displays the annotated image with masked obstructions.
3. **Binarization:** The annotated masks are batch-processed to produce standardized binary images: cloud regions are represented in white (pixel value 255), and the background is black (pixel value 0). These binary masks serve as crucial inputs for subsequent machine learning analyses. Fig. 1(d) depicts the corresponding binarized mask, where cloud regions appear in white and the background in black.

Each image in the dataset is further categorized according to the proportion of unobservable regions—areas obscured by heavy cloud cover or other obstructions that hinder the detection of celestial objects. Three categories are defined: High, Moderate, and Low Unobservable Regions, as summarized in Table 1. Notably, the Low Unobservable Region category includes a fixed 2% allocation due to the presence of buildings in certain images. In the adjusted 2,000-image dataset, Low Unobservable Region images account for 34.02%, Moderate for 29.77%, and High for 36.21%. This systematic categorization facilitates detailed analysis of the impact of cloud coverage on image processing and segmentation, thereby supporting the development and evaluation of robust nighttime sky analysis algorithms.

To provide intuitive insight into the classification scheme, Fig. 2 presents examples from each category. For each type, both the raw image and its corresponding mask are shown, along with the actual mask ratio

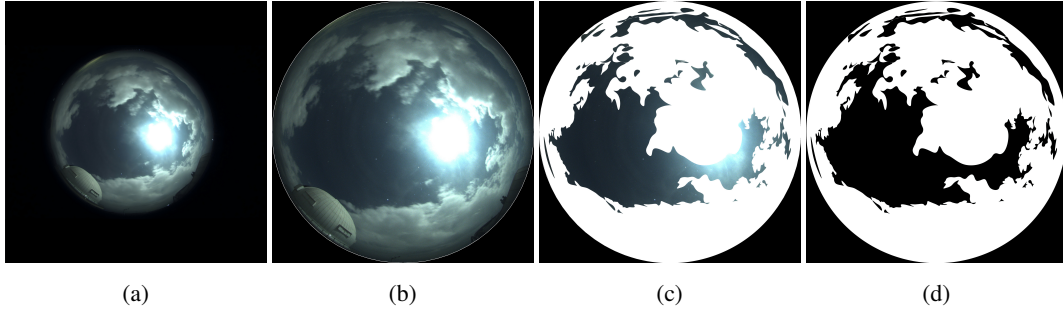


Fig. 1: Image preprocessing workflow: (a) Raw image captured by the all-sky camera; (b) Cropped and resized circular region; (c) Annotated image with masked obstructions; (d) Binarized mask highlighting cloud objects (white) and background (black).

Category	High Unobservable Region		Moderate Unobservable Region		Low Unobservable Region	
Raw Images						
Mask Images						
Mask Ratio	93.63%	100%	22.15%	21.47%	0.94%	1.04%

Fig. 2: Examples of full sky image and mask image in three categories. Two representative samples are shown for each category. The mask ratio is given at the bottom for each sample.

Table 1: Classification statistics of nighttime images captured by the all-sky camera based on the proportion of unobservable regions.

Statistic	High Unobservable Region	Moderate Unobservable Region	Low Unobservable Region
Images	724	596	680
Proportion	36.21%	29.77%	34.02%
Criterion	> 60%	2%–60%	< 2%

Notes: The statistics classify nighttime images based on the proportion of unobservable regions in each image. The columns correspond to the following categories: **High** (unobservable regions > 60%), **Moderate** (2%–60%), and **Low** (< 2%). Representative examples for each category are shown in Fig. 2.

(i.e., the proportion of unobservable regions). These examples visually demonstrate the diversity of sky conditions and the effectiveness of mask annotation in quantifying unobservable regions.

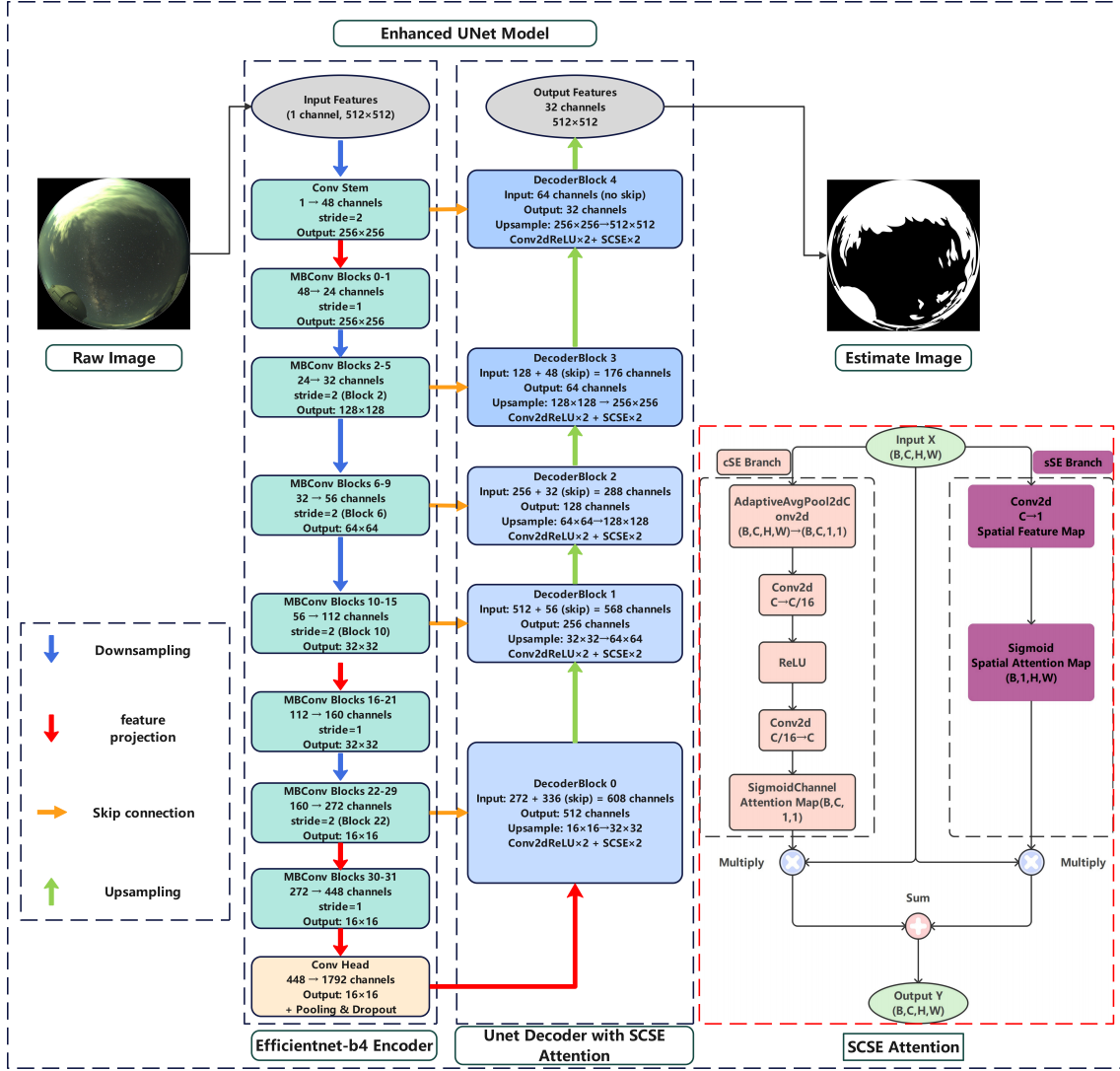


Fig. 3: The architecture of the proposed **Enhanced UNet**. The encoder is based on EfficientNet-B4, while the decoder incorporates SCSE attention modules for feature enhancement. Skip connections bridge the encoder and decoder to retain critical spatial details and contextual information. The output layer applies a Sigmoid activation for pixel-level precision.

3 UNOBSERVABLE REGION SEGMENTATION IN NIGHTTIME ALL-SKY IMAGES

3.1 Optimized Model Structure

This paper presents an improved model based on U-Net (Ronneberger et al. 2015), referred to as the **Enhanced UNet**. The architecture incorporates multiple enhancements over the baseline: an optimized encoder, an advanced decoder, effective regularization strategies, and a compound loss function to boost segmentation accuracy and model generalization.

The encoder in the Enhanced UNet employs a pre-trained EfficientNet-B4 (Tan & Le 2019) as the feature extractor. By leveraging ImageNet-pretrained weights, the encoder can extract features more efficiently, improve classification accuracy, and significantly reduce training time, ensuring the model learns diverse and meaningful representations from the input data.

In the decoder, the model incorporates SCSE attention modules (Hu et al. 2017), which dynamically recalibrate channel and spatial feature weights, enabling the network to focus more on salient information. A custom attention block can be incorporated if finer multi-scale feature extraction is needed. To further boost generalization and prevent overfitting, a Dropout layer is applied after the decoder’s final output. Skip connections throughout the network ensure both global and local contextual information are preserved, which improves pixel-level segmentation precision. The output layer utilizes a single-channel Sigmoid activation to generate pixel-wise probability maps.

To address the specific challenges of segmenting unobservable sky regions, the Enhanced UNet is trained using a **combined loss function** (Jadon 2020), which is a weighted sum of Binary Cross-Entropy (BCE) loss, Dice loss, and Intersection over Union (IoU) loss. This approach leverages the strengths of each loss—BCE for pixel-wise classification, Dice for class imbalance, and IoU for region-level overlap—to optimize different facets of the segmentation task. The combined loss is formulated as:

$$\text{CombinedLoss} = w_1 \cdot \text{loss}_{\text{BCE}} + w_2 \cdot \text{loss}_{\text{Dice}} + w_3 \cdot \text{loss}_{\text{IoU}},$$

where w_1 , w_2 and w_3 are tunable hyperparameters.

The overall architecture, as shown in Fig. 3, is tailored to maximize segmentation accuracy and robust feature representation, especially under complex nighttime sky conditions.

3.2 Experimental Configuration

The experimental setup employed in this study is detailed in Table 2. The Enhanced UNet model is trained and evaluated on a dataset comprising 2,000 high-resolution nighttime all-sky images. Each image is an RGB color image with a resolution of 2568×2538 pixels. Ground truth segmentation masks are manually annotated using image editing software, resulting in binary (black and white) masks precisely aligned with the original images. For model training, all input images are converted to grayscale, while the corresponding binary masks served as segmentation labels. The dataset is randomly divided into a training set (90%, 1,800 images) and a test set (10%, 200 images), ensuring both robust model learning and unbiased performance evaluation.

The network architecture is based on a pretrained EfficientNet-B4 encoder and a U-Net decoder enhanced with spatial and channel squeeze-and-excitation (SCSE) attention mechanisms. Model optimization is performed using the RAdam optimizer, with an initial learning rate of 2×10^{-4} and a weight decay of 1×10^{-5} . A cosine annealing warm restarts scheduler ($T_0 = 10$, $T_{\text{mult}} = 2$, minimum learning rate 1×10^{-6}) is employed to dynamically adjust the learning rate throughout training. The model is trained for up to 150 epochs with a batch size of 16, and early stopping is triggered if validation performance does not improve for 20 consecutive epochs. This experimental setup provides a robust and reproducible basis for subsequent model training and evaluation.

To objectively select the optimal loss weights, we performed a systematic grid search across the weight vector (w_1, w_2, w_3) with a step size of 0.1, enforcing the constraint $w_1 + w_2 + w_3 = 1$, $w_i \geq 0$. Each configuration was trained with identical data splits, optimization hyperparameters, and data augmentations, ensuring a fair and unbiased comparison. The validation set IoU served as the primary selection metric, with F1 score as a tiebreaker. Early stopping based on validation IoU (patience = 20 epochs) was applied

Table 2: Experimental Environment

Device	Configuration
Operating System	Windows 11 Chinese Home Edition
Processor	Intel(R) Core(TM) i9-14900K (3.20 GHz, 24 cores, 32 threads)
RAM	64 GB
PyTorch Version	2.5.1
GPU	NVIDIA GeForce RTX 4090 D
CUDA Version	11.8

Table 3: Comparison of the U-Net baseline and Enhanced U-Net on the expanded dataset. Loss parameters (w_1 , w_2 , w_3) and evaluation metrics (IoU, Precision, Recall, F1 Score) are shown. All Enhanced U-Net settings outperform the baseline; the best values are highlighted in bold.

Model	w_1	w_2	w_3	IoU	Precision	Recall	F1 Score
U-Net Baseline	-	-	-	0.9025	0.9401	0.9441	0.9320
Enhanced U-Net	0.1	0.3	0.6	0.9164	0.9531	0.9572	0.9468
	0.3	0.3	0.4	0.9178	0.9544	0.9598	0.9506
	0.5	0.3	0.2	0.9162	0.9527	0.9575	0.9465
	0.2	0.5	0.3	0.9159	0.9523	0.9569	0.9460
	0.4	0.4	0.2	0.9160	0.9525	0.9571	0.9462
	0.4	0.2	0.4	0.9212	0.9564	0.9602	0.9537
	1.0	0.0	0.0	0.9141	0.9509	0.9555	0.9443
	0.0	1.0	0.0	0.9146	0.9512	0.9558	0.9450
	0.0	0.0	1.0	0.9151	0.9519	0.9562	0.9457

to prevent overfitting. Each setting was evaluated over three random initializations, and we report the mean and standard deviation for all metrics (see Table 3 and Fig. 4). To verify that the optimum is not due to grid discretization, a finer search (step = 0.05) was conducted around the best weight, confirming ranking stability. The hold-out test set, which was never used during training or model selection, was used for final model assessment.

3.3 Experimental Results

Table 3 presents a comprehensive comparison between the U-Net baseline and several Enhanced UNet configurations with different loss weight settings (w_1 , w_2 , w_3). The evaluation metrics—Intersection over Union (IoU), Precision, Recall, and F1 Score—collectively reveal the segmentation performance and robustness of each model. The U-Net baseline serves as a reference, while all Enhanced UNet variants outperform it across all metrics, demonstrating the effectiveness of the proposed architectural and loss function improvements.

Among the Enhanced UNet configurations, the combination $w_1 = 0.4$, $w_2 = 0.2$, $w_3 = 0.4$ delivers the best results, achieving an IoU of 0.9212, Precision of 0.9564, Recall of 0.9602, and F1 Score of 0.9537. These improvements are especially significant for IoU and F1 Score, which are critical for accurate night-

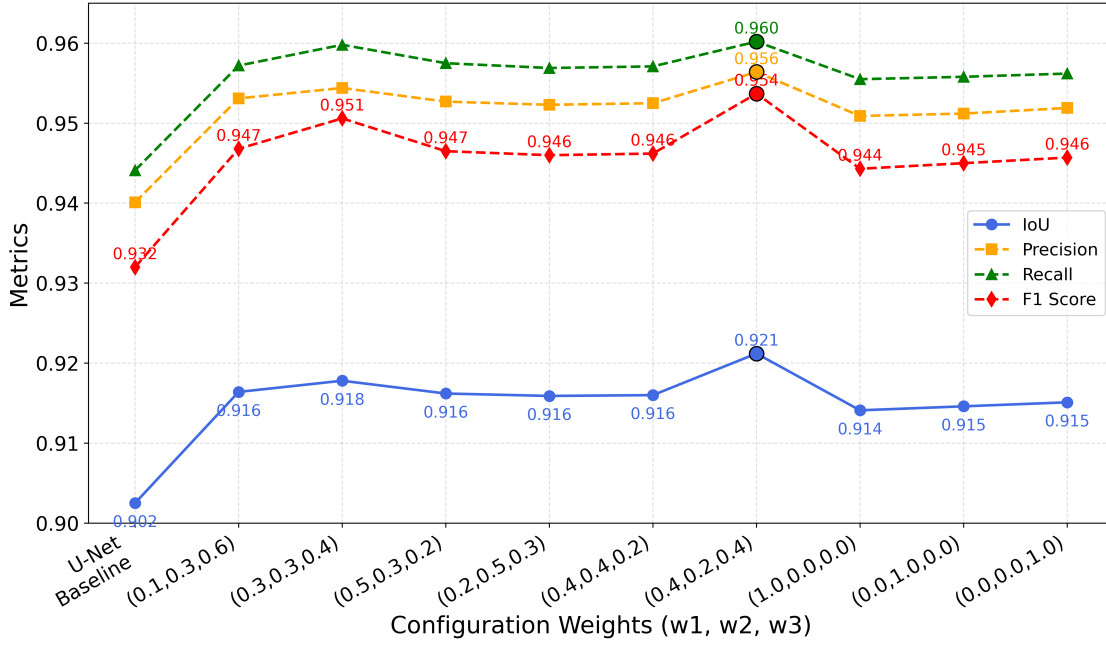


Fig. 4: Performance evaluation of the U-Net baseline and Enhanced UNet configurations on the expanded dataset. The Enhanced UNet achieves consistently better IoU, Precision, Recall, and F1 Score. The configuration ($w_1 = 0.4, w_2 = 0.2, w_3 = 0.4$) yields the best results.

time sky segmentation. Experiments with a single loss component ($w_1 = 1.0, w_2 = 0.0, w_3 = 0.0$, etc.) yield lower performance, underlining the importance of a balanced hybrid loss. Fig. 4 clearly visualizes these trends, highlighting the stable and substantial performance gains achieved with balanced loss settings.

Beyond quantitative metrics, Fig. 5 offers qualitative comparisons in three representative unobservable region scenarios (high, moderate, and low). For each, the raw input image, predicted mask, ground truth, overlay, and difference map are provided. The final row reports the pixel-wise difference ratio, directly reflecting segmentation error. Overlay images exhibit close alignment between predictions and ground truth, while the difference maps (red for false positives, blue for false negatives) provide insight into model errors. Notably, the Enhanced UNet achieves a minimal difference ratio of just 0.33% in the low unobservable region case, and maintains high accuracy in more challenging scenarios.

The experiments are conducted on a robust dataset of 1,800 training and 200 test images, ensuring statistical reliability. Compared to the baseline, the Enhanced UNet demonstrates substantial improvements—particularly for IoU and F1 Score—in both quantitative and qualitative evaluations. These results validate the effectiveness of the SCSE attention mechanism and hybrid loss in addressing nighttime sky segmentation challenges.

In summary, the hybrid loss design consistently delivers superior segmentation performance and remarkable robustness. Intuitively, BCE loss ensures stable pixel-wise learning and enhances boundary precision; Dice loss addresses class imbalance prevalent in large unobservable regions; and IoU loss reinforces region-level consistency and connectivity. The synergy of these components enables the network to generalize effectively across diverse nighttime conditions. Ablation studies further confirm that hybridizing these losses is essential, as single-loss or pairwise combinations result in noticeably reduced accuracy.

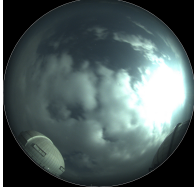
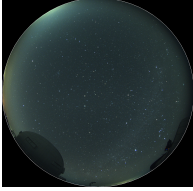
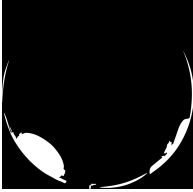
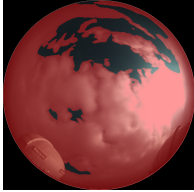
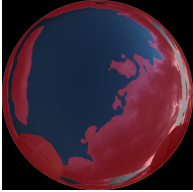
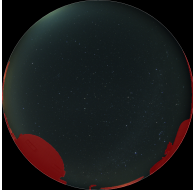
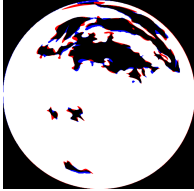
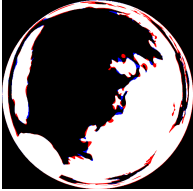
Category	High Unobservable Region	Moderate Unobservable Region	Low Unobservable Region
Raw Image			
Model Prediction			
Ground Truth			
Overlay Image			
Difference Map			
Difference Ratio	3.56%	3.96%	0.33%

Fig. 5: Visualization of model performance for three categories (High, Moderate, and Low Unobservable Region). From top to bottom, the panels show: the raw sky image; the model prediction mask; the ground truth mask; the overlay image (showing the overlap of model prediction and ground truth on the raw image); and the difference map. The difference map highlights discrepancies between the model prediction and ground truth: red indicates areas incorrectly predicted as unobservable regions by the model (false positives), while blue indicates unobservable regions missed by the model (false negatives). The final row Difference Ratio shows the pixel-wise discrepancy between the model prediction and ground truth.

4 APPLICATIONS OF CLOUD COVERAGE DETECTION IN THE UNATTENDED OPERATION OF MEPHISTO TELESCOPE

In the previous section, the Enhanced UNet model is employed to identify unobservable regions. To incorporate the results into the Observation Control System (OCS), the pointing coordinates (RA, DEC) of the telescope during observation needs to convert to the pixel coordinates (x, y) on the all-sky camera image.

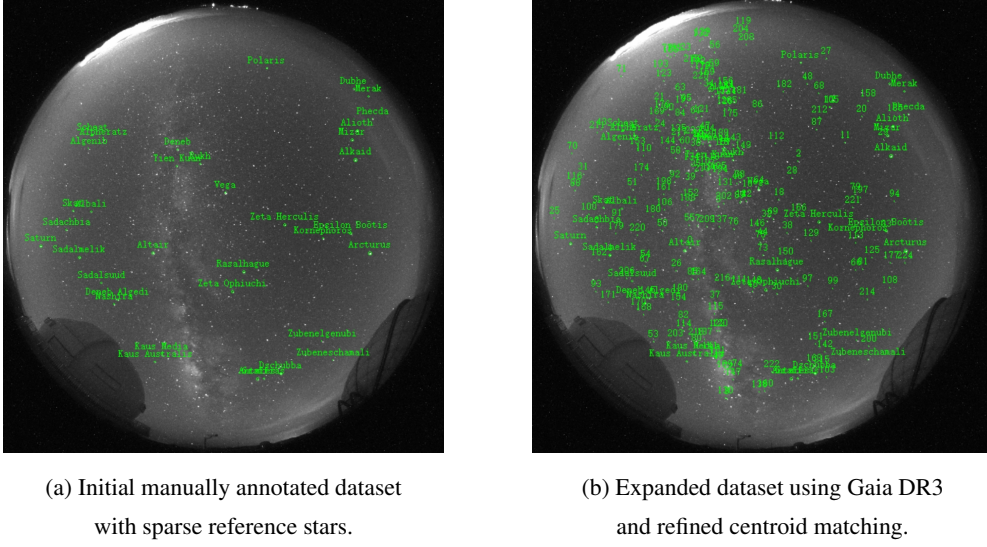


Fig. 6: Two-stage dataset construction for all-sky coordinate mapping. (a) An initial sparse WCS is built from manually selected stars spanning zenith to horizon. (b) A refined, dense calibration dataset is created by matching Gaia DR3 stars with local centroids, allowing high-precision model fitting.

First, this section will introduce the modeling of the all-sky camera and its application in the telescope. Then we will establish a transformation system to link between telescope pointing coordinates and image pixel coordinates.

4.1 Modeling of World Coordinates and Image Coordinates for All-sky Camera Image

Accurately determining whether a telescope’s pointing intersects unobservable regions, and achieving a precise transformation between celestial coordinates and image pixels, are essential yet nontrivial tasks in wide-field astronomical observations. Due to the geometric distortion inherent in all-sky camera systems, standard astrometric tools such as `astrometry.net` (Lang et al. 2010) are inadequate for directly deriving a reliable mapping from equatorial coordinates (RA, Dec) to the pixel coordinates (x, y) . Here, we adopt a two-stage transformation approach. First, equatorial coordinates are converted to horizontal coordinates—namely, altitude (Alt) and azimuth (Az)—based on the time and geographical location of observatory. Then, a Zenithal Equal-Area (ZEA) projection (Calabretta & Greisen 2002) is employed to project the (Alt, Az) pair onto the pixel coordinates. Compared to traditional tangent-plane (TAN) projections, the ZEA projection offers better performance in preserving areal relationships and handling the wide angular extent from zenith to horizon, making it particularly well-suited for all-sky imaging applications (Xie et al. 2025).

4.1.1 Database Building

To establish the mapping between horizontal coordinates and image pixel coordinates, we adopted a two-stage dataset construction strategy. As shown in Fig. 6a, we first selected a high-quality all-sky image captured under clear, moonless conditions and manually annotated a small number of bright stars from zenith to horizon. This initial sparse dataset allowed us to derive a coarse World Coordinate System (WCS) solution using the ZEA projection model.

Based on the initial calibration calibration, we retrieved all Gaia DR3 stars brighter than magnitude 8 within $\pm 90^\circ$ zenith distance and $\text{pmra} < 20$ and $\text{pmdec} < 20$. Their (Alt, Az) coordinates are projected into preliminary (x, y) image coordinates. We then used `Source Extractor` (Bertin & Arnouts 1996) to identify actual star positions in the image and applied nearest-neighbor or local optimization matching to refine the associations. The resulting dense set of matched star pairs, illustrated in Fig. 6b, enabled robust polynomial fitting and significantly improved model accuracy.

4.1.2 Coordinate Transformation Model

The coordinate transformation between celestial coordinates and pixel coordinates is realized through a two-stage approach: (1) projection from horizontal coordinates (Alt, Az) to intermediate planar coordinates (ξ, η) via ZEA projection, and (2) polynomial mapping from (ξ, η) to image pixel coordinates (x, y) .

The ZEA projection centers on the zenith and is well-suited for all-sky cameras with ultra-wide fields of view due to its area-preserving property. Given the Alt and Az of a celestial object, the zenith distance c is computed as:

$$c = 90^\circ - \text{Alt} \quad (1)$$

Converting to radians:

$$c_{\text{rad}} = \frac{\pi}{180} \cdot c, \quad \text{az}_{\text{rad}} = \frac{\pi}{180} \cdot \text{Az} \quad (2)$$

The ZEA projection equations are:

$$\begin{aligned} \xi &= 2R \sin\left(\frac{c_{\text{rad}}}{2}\right) \cos(\text{az}_{\text{rad}}) \\ \eta &= 2R \sin\left(\frac{c_{\text{rad}}}{2}\right) \sin(\text{az}_{\text{rad}}) \end{aligned} \quad (3)$$

Here, R is the projection scale factor, typically set such that the projected radius matches the pixel radius from zenith to horizon, ensuring that (ξ, η) are in pixel units.

To account for optical and projection distortions, we fit the intermediate coordinates (ξ, η) to the final image coordinates (x, y) using a two-dimensional polynomial of order n :

$$\begin{aligned} x &= \sum_{i+j \leq n} a_{ij} \xi^i \eta^j, \\ y &= \sum_{i+j \leq n} b_{ij} \xi^i \eta^j \end{aligned} \quad (4)$$

For $n = 2$, this yields a total of 6 terms for x and 6 terms for y , resulting in 12 coefficients to be determined. These include constant, linear, and quadratic components, enabling the model to capture nonlinear distortions introduced by the camera optics and the projection process.

4.1.3 Residual Analysis

To evaluate the accuracy of the coordinate transformation model, we calculate the residual error for each calibration star as the Euclidean distance between its predicted and measured image coordinates:

$$\text{residual} = \sqrt{(x_{\text{pred}} - x_{\text{meas}})^2 + (y_{\text{pred}} - y_{\text{meas}})^2} \quad (5)$$

where $(x_{\text{meas}}, y_{\text{meas}})$ are the centroid positions extracted from the all-sky image, and $(x_{\text{pred}}, y_{\text{pred}})$ are the corresponding predicted positions from the transformation model.

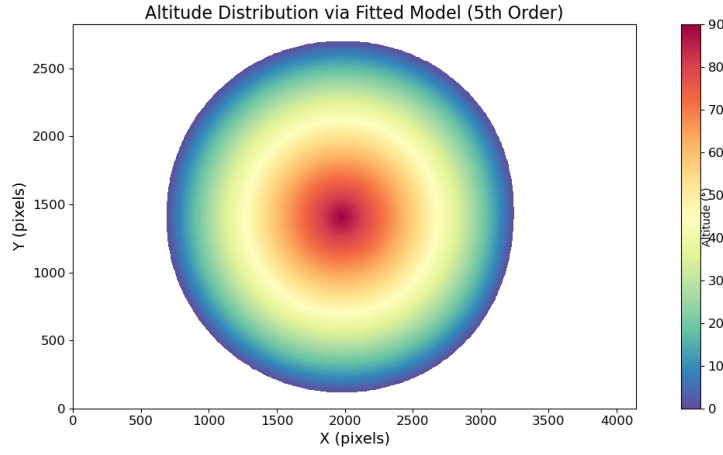
The mean residual across all matched stars quantifies the overall accuracy of the geometric model. Smaller residuals indicate improved correspondence between the projected sky model and actual stellar positions in the image.

4.1.4 Model Fitting and Discussion

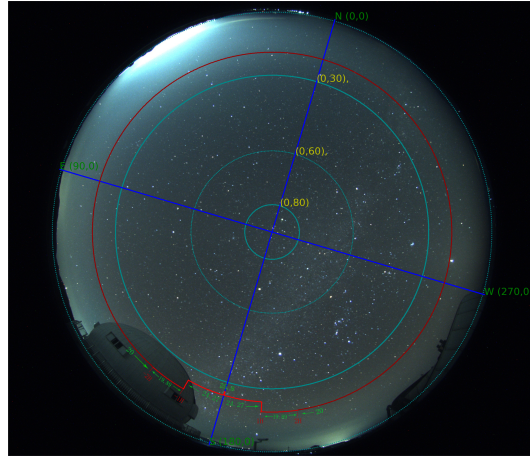
To model the celestial-to-pixel coordinate transformation, we implement a fifth-order polynomial ($n = 5$) for the transformation described in Eq. (4). This yields a mean residual of 0.95 pixels, corresponding to an image scale of 0.0621 degrees per pixel. To evaluate the quality of the transformation, we generate an ALT distribution map across the image plane, where pixel colors represent modeled ALT angles. As shown in Fig. 7, the top panel illustrates the fitted ALT field. The resulting smooth and radially symmetric gradient confirms that the polynomial mapping preserves the sky’s geometric structure. The bottom panel presents a real all-sky image under clear sky conditions, overlaid with the calculated ALT contours and AZ reference lines derived from the transformation. Concentric circles correspond to ALT levels (e.g., 30° , 60° , and 80°), while cardinal directions (N, E, S, W) are labeled in green. A red curve indicates the practical ALT limit of the Mephisto telescope, which defines the lower boundary for reliable observations due to mechanical constraints. The excellent agreement between the modeled ALT contours and the observed stellar background demonstrates the effectiveness and low distortion of the fitted projection model, validating its use in visibility estimation and dynamic scheduling.

Building upon the Enhanced UNet segmentation model (Section 3) that identifies unobservable regions and the coordinate transformation pipeline (Section 4) using ZEA projection to map celestial coordinates to image pixels, this system establishes a coordinate mapping that transcends geometric alignment. It enables real-time, pixel-level visibility queries for any telescope pointing (defined by RA/DEC): by converting the pointing to its corresponding location in the all-sky image and ultimately the ALT-AZ system, the system instantly determines if it intersects a segmented obscured area (e.g., cloud cover). Furthermore, a configurable altitude buffer (e.g., $5\text{--}10^\circ$) prevents scheduling targets near the image edge – where observing conditions degrade rapidly – significantly increasing scheduler resilience.

The system’s capability goes beyond binary visibility judgment (whether the target is observable). The output of the segmentation model supports quantitative evaluation of sky quality: by calculating the proportion of unobservable pixels relative to the total projected area of the celestial dome (whether caused by clouds, moonlight scattering, or other atmospheric effects), cloud coverage can be obtained in real-time as a percentage. This type of statistics is crucial for scheduling systems, as it can be used to prioritize clear sky areas, reweight target queues, or evaluate the feasibility of observations throughout the night. Although moonlight scattering may sometimes exaggerate cloud coverage estimates, this “pollution” itself corresponds to actual decreasing photometric degradation, and therefore has guiding significance for sky survey operations. This capability also helps alleviate the limitations of infrared cloud sensors (IRCS), which may provide misleading readings when encountering high humidity or moonlight during unmanned operation. This function forms the basis for visibility aware scheduling. It will be implemented into the Mephisto Observation Control System (OCS) in future. An example demonstrating how the segmentation and coordinate mapping pipeline supports cloud-aware scheduling is shown in Fig. 8. Once deployed, OCS



(a) Modeled ALT distribution in image coordinates.



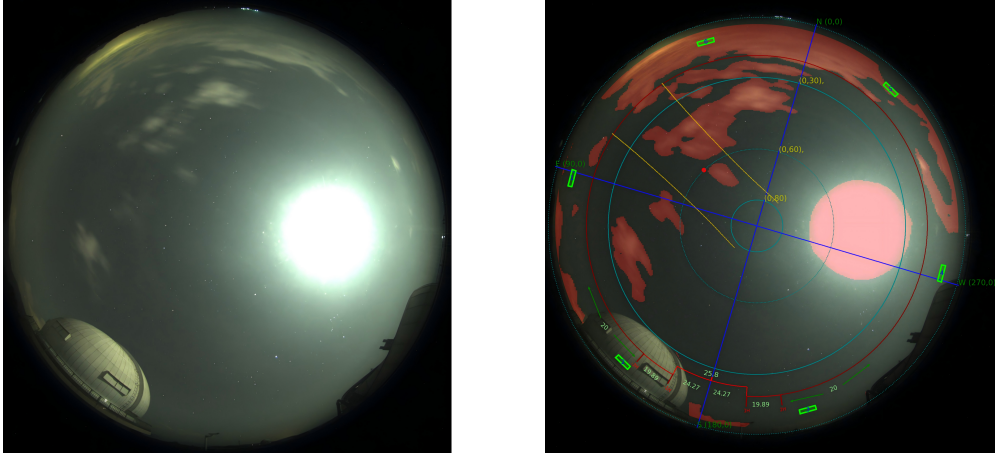
(b) Actual all-sky image with ALT/AZ overlay and pointing limit.

Fig. 7: Coordinate transformation evaluation using ZEA projection. (a) Symmetric ALT gradient centered on zenith; (b) Geometric alignment between model and stellar positions. The red curve marks the lower limit of the Mephisto telescope pointing.

Table 4: Comparison of Nighttime Sky Visibility estimation Methods

Feature	Photometry/Extinction	Photometry-Free (PFSAVE)	Segmentation-based (This work)
Input	FITS + catalog	FITS + catalog	All-sky images
Granularity	Grid ($1-2^\circ$)	Grid ($1-2^\circ$)	Pixel-level ($<0.1^\circ$)
Latency	Seconds–minutes	Seconds	Milliseconds
Moon/Cloud Robust	Poor	Moderate	good
Interpretability	Physical extinction (mag)	Limiting mag per grid	Pixel mask, supports aggregation
Accuracy	~ 0.05 mag (clear)	$r > 0.85$ (vs. expert)	IoU 0.92, F1 0.95 (pixel)
Primary Application	Sky condition statistics	Dynamic grid scheduling	Real-time fine-grained scheduling

will automatically compare telescope pointing with segmentation mask, avoid unobservable areas (or close to an angular distance buffer around it), and dynamically rearrange the observation.



(a) Raw all-sky image captured at Lijiang Observatory under strong moonlight conditions.

(b) Segmentation mask (red) indicating cloud/moon-obscured areas.

Fig. 8: Example of cloud-aware scheduling based on segmentation and coordinate mapping. The telescope's pointing (red dot) lies near a cloud boundary.

We compare our method with others. Photometry/extinction-based methods (Zhi et al. 2024) rely on extracting stellar sources and comparing their measured fluxes to catalog values within spatial grids (such as HEALPix). This framework provides physically interpretable extinction and transparency estimates, and its results are often benchmarked by quantitative agreement with manual ratings (e.g., median error of ~ 0.05 mag in clear conditions). However, these methods are sensitive to the quality of source extraction and accurate catalog matching. Performance is readily influenced by strong moonlight, sky gradients, incomplete calibration, and photometric outliers. Their spatial granularity is set by the chosen grid size (e.g., typically 1° – 2° per cell), and the computation introduces non-negligible latency due to repeated photometric analysis, which may not meet the requirements of low-latency, real-time scheduling for robotic observations.

Photometry-free algorithms, such as the PFSAVE method (Xie et al. 2025), infer visibility in each grid cell based not on photometric calibration, but on the limiting magnitude of robustly detected stars. This approach incorporates precise astrometric transformations and carefully designed detection metrics (e.g., S/N thresholds, T-metrics) to minimize false positives from noise, clouds, or moonlight. Although PFSAVE can nominally process lower-quality images, high-precision visibility estimation still requires FITS data and reliable star matching. The resulting grid-wise visibility maps correlate well with expert assessments (e.g., Spearman $r > 0.85$ under cloud cover). However, its spatial resolution is fundamentally limited by the grid definition, and real-time performance is affected by the computational demands of star extraction and assignment.

In contrast, our segmentation-based approach adopts an end-to-end deep learning framework. It directly produces pixel-level visibility masks from standard all-sky images, eliminating the need for specific FITS photometry, stellar extraction, or external catalogs. This enables real-time, high-resolution sky condition mapping robust to faint clouds, cloud edges, and complex illumination, including scattered moonlight. Empirical experiments demonstrate that our method achieves millisecond-level inference per frame on modern GPUs, supporting true online operation in unattended environments. Furthermore, segmentation masks can be aggregated into arbitrary grid scales if needed, enabling both fine-grained decision-making and

seamless integration with grid-based pipelines. The main properties and performance of the three methods are presented in Table 4.

5 CONCLUSION

In this study, we have developed a deep learning-based framework for the segmentation and mapping of unobservable regions—such as clouds and moon-scattered areas—in nighttime all-sky images, with the goal of enhancing the efficiency and robustness of autonomous astronomical observations. Specifically, we designed an enhanced UNet architecture that incorporates an EfficientNet-B4 encoder, SCSE attention modules in the decoder, and a compound loss function combining Binary Cross-Entropy, Dice, and IoU losses. Extensive experiments conducted on a manually annotated dataset of 2,000 nighttime all-sky images demonstrate that our model achieves state-of-the-art segmentation performance. With the optimal loss weight configuration ($w_1 = 0.4$, $w_2 = 0.2$, $w_3 = 0.4$), the model attains an Intersection over Union (IoU) of 0.9212, Precision of 0.9564, Recall of 0.9602, and F1 Score of 0.9537 on the test set. Qualitative analysis further reveals that the pixel-wise difference between predicted and ground-truth masks is as low as 0.33% under clear-sky conditions, and remains below 4% even in challenging cases with substantial cloud coverage.

To enable practical deployment in robotic observatories, we have established a Zenithal Equal Area (ZEA)-based coordinate transformation pipeline, which accurately maps telescope pointing coordinates (RA, Dec) to all-sky image pixels. The transformation model, fitted with a fifth-order polynomial, achieves a mean residual of just 0.95 pixels, ensuring highly precise alignment between celestial and image coordinates. This system supports real-time, pixel-level visibility checks for any telescope pointing, and provides quantitative cloud coverage statistics to inform dynamic scheduling decisions. The framework is planned to be integrated into the Mephisto telescope’s observation control system (OCS), thereby enabling cloud-aware, fully autonomous sky survey operations.

Looking ahead, several directions remain for further improvement. To better capture dynamic weather patterns, the development of a temporal model for cloud evolution could enable short-term predictions of sky conditions. In addition, further refinement of the segmentation model will be pursued to address photometrically challenging scenarios, such as extremely faint clouds or strong moonlight contamination. Seamless integration of the segmentation and coordinate mapping pipeline with the observatory control system will also be a focus of future work, supporting real-time, weather-adaptive scheduling and robust data quality control. Finally, expanding the annotated dataset and exploring transfer learning strategies may further enhance the generalization capability of the framework to other observatories and diverse atmospheric environments.

In summary, the developed deep learning framework enables precise segmentation of unobservable sky regions and accurate mapping between celestial and image coordinates, greatly enhancing the autonomy and adaptability of astronomical observations. With its planned integration into the Mephisto observation control system, the method lays a solid foundation for real-time, weather-aware sky surveys. Looking forward, expanding the dataset and introducing temporal modeling will further broaden the applicability and robustness of this framework across diverse observational environments.

Acknowledgements This work is supported by the National Key R&D Program of China (Grant No. 2024YFA1611600), the Yunnan Fundamental Research Projects (Grant No. 202501AT070455), and the Yunnan Revitalization Talent Support Program, including the Science & Technology Champion Project (Grant No. 202005AB160002), the Innovation Team Project (Grant No. 202105AE160021), and the Top Team Project (Grant No. 202305AT350002). The authors also acknowledge support from the Key Laboratory of Survey Science of Yunnan Province with Project No. 202449CE340002. We would like to thank the staff of the Lijiang Observatory for their assistance with data acquisition and system maintenance. Additionally, we are grateful to the members of the Mephisto project team for valuable discussions and technical assistance. Special thanks are extended to the anonymous reviewers for the constructive comments and suggestions, which have greatly improved the quality of this manuscript. We also acknowledge the use of the Gaia DR3 catalog and the open-source software tools utilized in this research.

References

- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393 12
- Calabretta, M. R., & Greisen, E. W. 2002, *A&A*, 395, 1077 11
- Calbó, J., & Sabburg, J. 2008, *Journal of Atmospheric and Oceanic Technology*, 25, 3 2
- Cazorla, A., Olmo, F. J., & Alados-Arboledas, L. 2008, *Journal of the Optical Society of America A*, 25, 29 2
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, *arXiv e-prints*, arXiv:1612.05560 2
- Cheng, H.-Y., & Lin, C.-L. 2017, *Atmospheric Measurement Techniques*, 10, 199 3
- Cordier, B., Götz, D., Motch, C., & SVOM Collaboration. 2018, *Mem. Soc. Astron. Italiana*, 89, 266 2
- Dekany, R., Smith, R. M., Riddle, R., et al. 2020, *PASP*, 132, 038001 2
- Du, G., Cheng, Y., Yang, Y.-P., et al. 2025, *ApJ*, 989, 19 2
- Duev, D. A., Mahabal, A., Masci, F. J., et al. 2019, *MNRAS*, 489, 3582 3
- Fabel, Y., Nouri, B., Wilbert, S., et al. 2022, *Atmospheric Measurement Techniques*, 15, 797 3
- Forster, F. 2021, in *American Astronomical Society Meeting Abstracts*, Vol. 237, American Astronomical Society Meeting Abstracts, 423.03 3
- Fu, M.-X., Song, Y., Lv, J.-M., et al. 2024, *Chinese Physics C*, 48, 095001 3
- Ghonima, M. S., Urquhart, B., Chow, C. W., et al. 2012, *Atmospheric Measurement Techniques*, 5, 2881 2
- Hausen, R., & Robertson, B. E. 2020, *ApJS*, 248, 20 3
- Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. 2017, *arXiv e-prints*, arXiv:1709.01507 7
- Jadon, S. 2020, *arXiv e-prints*, arXiv:2006.14822 7
- Jia, P., Liu, W., Liu, Y., & Pan, H. 2023, *ApJS*, 264, 43 3
- Kazantzidis, A., Tzoumanikas, P., Bais, A. F., Fotopoulos, S., & Economou, G. 2012, *Atmospheric Research*, 113, 80 2
- Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2010, *AJ*, 139, 1782 11
- Lewis, P. M., Rogers, H., & Schindler, R. H. 2010, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7735, *Ground-based and Airborne Instrumentation for Astronomy III*, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 77353C 2

- Li, Q., Lu, W., Yang, J., & Wang, J. Z. 2012, *IEEE Geoscience and Remote Sensing Letters*, 9, 417 3
- Liu, S., Zhang, L., Zhang, Z., Wang, C., & Xiao, B. 2015, *IEEE Geoscience and Remote Sensing Letters*, 12, 354 3
- Liu, X., Chen, X., Zhang, Y., et al. 2025, *GRB Coordinates Network*, 40197, 1 2
- Muthukrishna, D., Narayan, G., Mandel, K. S., Biswas, R., & Hložek, R. 2019, *PASP*, 131, 118002 3
- Ofek, E. O., Laher, R., Law, N., et al. 2012, *PASP*, 124, 62 2
- Redman, B. J., Shaw, J. A., Nugent, P. W., Clark, R. T., & Piazzolla, S. 2018, *Optics Express*, 26, 11276 2
- Ronneberger, O., Fischer, P., & Brox, T. 2015, *arXiv e-prints*, arXiv:1505.04597 6
- Shields, J. E., Karr, M. E., Johnson, R. W., & Burden, A. R. 2013, *Appl. Opt.*, 52, 1605 2
- Souza-Echer, M. P., Pereira, E. B., Bins, L. S., & Andrade, M. A. R. 2006, *Journal of Atmospheric and Oceanic Technology*, 23, 437 2
- Sun, R., Jia, P., Sun, Y., et al. 2023, *AJ*, 166, 235 3
- Takato, N., Okada, N., Kosugi, G., et al. 2003, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4837, *Large Ground-based Telescopes*, ed. J. M. Oschmann & L. M. Stepp, 872 2
- Tan, H., Cen, X.-F., Qian, T.-L., & Wang, J.-C. 2002, *Bulletin of the Astronomical Society of India*, 30, 881 3
- Tan, M., & Le, Q. V. 2019, *arXiv e-prints*, arXiv:1905.11946 6
- Taravat, A., Del Frate, F., Cornaro, C., & Vergari, S. 2015, *IEEE Geoscience and Remote Sensing Letters*, 12, 666 2
- Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, *PASP*, 130, 064505 2
- Wang, C.-J., Bai, J.-M., Fan, Y.-F., et al. 2019, *Research in Astronomy and Astrophysics*, 19, 149 3
- Wang, T., Liu, G., Cai, Z., et al. 2023, *Science China Physics, Mechanics, and Astronomy*, 66, 109512 2
- Xie, H., Zhi, H., Kang, Z., et al. 2025, *Astronomical Techniques and Instruments*, 2, 52 11, 15
- Xie, W., Liu, D., Yang, M., et al. 2020, *Atmospheric Measurement Techniques*, 13, 1953 3
- Xin, Y.-X., Bai, J.-M., Lun, B.-L., et al. 2020, *Research in Astronomy and Astrophysics*, 20, 149 2, 4
- Yin, Y.-H. I., Fang, Y., Zhang, B.-B., et al. 2025, *ApJ*, 989, L39 2
- Yuan, W., Dai, L., Feng, H., et al. 2025, *arXiv e-prints*, arXiv:2501.07362 2
- Yuan, X., Li, Z., Liu, X., et al. 2020, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 11445, *Ground-based and Airborne Telescopes VIII*, ed. H. K. Marshall, J. Spyromilio, & T. Usuda, 114457M 2, 3
- Zhang, J., Lian, J., Kumar, B., et al. 2025, *The Astronomer's Telegram*, 17022, 1 2
- Zhi, H., Wang, J., Zhang, X., et al. 2024, *PASP*, 136, 035002 15