Supplementary information for IMPMCT:a dataset of Integrated Multi-source Polar Meso-Cyclone Tracks

Runzhuo Fang[1,2], Jinfeng Ding[1,2], Wenjuan Gao[1,2], Xi Liang[3], Zhuoqi Chen[4], Chuanfeng Zhao[5], Haijin Dai[1,2], Lei Liu[1,2]

[1]College of Meteorology and Oceanography, National University of Defense Technology, Changsha, China
[2]Key Laboratory of High Impact Weather(special), China Meteorological Administration
[3]Key Laboratory of Marine Hazards Forecasting, National Marine Environmental Forecasting Center, Ministry of Natural Resources, Beijing, China
[4]School of Geospatial Engineering and Science, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-sen University, Zhuhai, China
[5]Department of Atmospheric and Oceanic Sciences, School of Physics, and China Meteorological Administration Tornado Key Laboratory, Peking University, Beijing, China

*Correspondence to: Jinfeng Ding (*dingjinfeng@nudt.edu.cn*)*

Table S1 | performance of YOLOv8-obb-pose model

We specifically provide the YOLOv8-obb-pose model weights to enable other researchers to replicate the model, alongside a small validation dataset for performance evaluation and to facilitate its implementation (Fang, 2025). The validation dataset comprises 1334 Vortex-Centered Infrared (VCI) images from the Nordic Sea region spanning the years 2001 and 2023, with 500 cyclone-containing images of the dataset. None of these images are involved in any training process of the model. The model's performance is evaluated using three common metrics—precision, recall, and mean average precision (mAP)—for both keypoint detection and oriented bounding box prediction tasks on this validation set.

$$Precesion = \frac{TP}{TP + FP} = \frac{TP}{predictions} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{ground\ truths} \tag{2}$$

$$L_{oks} = exp^{\left(-\frac{d^2}{2s^2k^2}\right)} \tag{3}$$

In Eq. (1) and (2), for the oriented bounding box task, a true-positive (TP) is defined as a correctly predicted class when the predicted bounding box overlaps sufficiently with the ground-truth bounding box, with the overlap degree measured by the Intersection over Union $(IoU, Intersection/Union$ between the predicted and ground-truth boxes). For the keypoint detection task, a TP occurs when the predicted keypoint is sufficiently close to the ground-truth keypoint, with proximity quantified by the Object Keypoint Similarity (OKS) metric defined in Eq. (3) (Maji et al., 2022), in which d denotes the Euclidean distance between the predicted and true keypoint locations, k represents the keypoint importance weight, and s corresponds to the area of the object's bounding box. Additionally, false positives (FP) are instances where predicted bounding boxes (or keypoint locations) fail to meet the required IoU (or L_{oks}) threshold with any ground-truth object bounding. False negatives (FN) occur when ground-truth objects (bounding boxes/keypoints) are undetected or fail to meet the IoU/L_{oks} threshold with predictions.

The model's performance on this validation dataset with IoU (L_{oks}) threshold set as 0.5 and confidence threshold set as 0.3 is summarized in Table S1. The mAP50-95 metric represents the mean average precision across IoU or L_{oks} thresholds from 0.5 to 0.95, calculated by dynamically optimizing confidence thresholds to balance precision and recall at each threshold level. The model achieved notable precision and recall values in both prediction tasks. The mAP50-95 for bounding box prediction is significantly lower than mAP50, indicating that the model performes better on well-defined cloud feature samples compared to ambiguous ones. In contrast, the keypoint prediction for cyclone center locations demonstrates consistently high accuracy regardless of sample complexity, suggesting that boundary box prediction for cyclones is more challenging than localizing their centers. Overall, comma-shaped clouds exhibit significantly higher prediction accuracy than spiral clouds. This discrepancy may stem from class imbalance in the dataset or the model's incomplete ability to distinguish between spiral and comma-shaped cloud structures, implying substantial room for improvement in the model's generalization capabilities.

Table S1 The Yolov8-obb-pose model's performance on this validation dataset

Class	Instances	BOX(P	R	mAP50	mAP50 -95)	Pose(P	R	mAP50	mAP50 -95)
comma	285	0.94	0.87	0.92	0.59	0.97	0.90	0.94	0.94
spiral	215	0.83	0.88	0.85	0.45	0.88	0.96	0.93	0.93
all	500	0.88	0.88	0.88	0.52	0.92	0.93	0.94	0.94

References

Maji, D., Nagori, S., Mathew, M., and Poddar, D.: YOLO-pose: Enhancing YOLO for multi person pose estimation using object keypoint similarity loss, 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), New Orleans, LA, USA, 2636–2645, https://doi.org/10.1109/CVPRW56347.2022.00297, 2022.

Fang, R.: validation dataset for yolov8-obb-pose cyclone-detect-model, Zenodo[data set], https://doi.org/10.5281/zenodo.15119534, 2025