https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

Deep-learning-based stage-aware classification of Arctic melt ponds and their microtopographic associations from high-resolution UAV imagery

Wenyu Shen[1,2], Xiaoping Pang[1,2*], Meng Qu[3*], Guokun Lv[3], Xunxiang Wu[1,2], Farui Jiang[1,2], Jianbo Shao[1,2], Na Li[3]

1 Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430079, China

2 Key Laboratory of Polar Environmental Monitoring and Public Governance, Ministry of Education (Wuhan University), Wuhan 430079, China

3 Key Laboratory of Polar Science, Ministry of Nature Resources, Polar Research Institute of China, Shanghai 200136, China

***Corresponding author.** pxp@whu.edu.cn(X.Pang); qumeng@pric.org.cn(M.Qu)*

Abstract. Melt ponds are a key component of the summer Arctic sea-ice surface because their formation and evolution strongly affect surface albedo, energy absorption, meltwater redistribution, and sea-ice mass balance. Most previous remote-sensing studies have treated melt ponds as a single surface class, limiting the characterization of heterogeneous pond states during late-summer melt and refreezing. In this study, we developed MP-Unet, a stage-aware semantic segmentation framework for identifying Open, Transitional, and Frozen Melt Ponds from high-resolution unmanned aerial vehicle imagery acquired during the 14th Chinese National Arctic Research Expedition. MP-Unet integrates residual blocks with channel attention, atrous spatial pyramid pooling, attention-gated skip connections, and an auxiliary binary segmentation head. The full model achieved an F1-score of 0.9440 and a mean intersection over union of 0.7466, with class-specific IoU values of 0.5897, 0.7544, and 0.6539 for Open, Transitional, and Frozen Melt Ponds, respectively. Stage-resolved mapping revealed marked spatial heterogeneity among the five observation sites, while Transitional Melt Ponds accounted for approximately 80.3% of the total pond area in the pooled sample. Pond area–frequency distributions showed a general scale-dependent decline and a sparse large-area tail, although the strength of the fitted scaling relationship varied among sites. Object-level analysis further showed that Frozen Melt Ponds generally had more compact and regular shapes, whereas Open Melt Ponds exhibited broader circularity distributions extending toward lower values. DEM-assisted analysis indicated significant stage-

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

dependent differences in local relative elevation: Transitional Melt Ponds occupied lower local topographic positions than Frozen Melt Ponds, despite the absence of significant differences in distance to the nearest ridge-like feature. These findings demonstrate that stage-aware classification provides information beyond conventional binary melt pond mapping by linking surface-state identification with pond morphology and local microtopographic position. The proposed framework offers a practical basis for fine-scale observations of Arctic sea-ice surface evolution and for the validation and improvement of satellite and numerical melt pond products.

Key words: Arctic sea ice; melt ponds; high-resolution UAV imagery; semantic segmentation; stage-aware classification; microtopography

Highlights. A stage-aware MP-Unet was developed for four-class segmentation of Arctic melt ponds from high-resolution UAV imagery. The proposed model achieved an mIoU of 0.7466 and an F1-score of 0.9440. Transitional Melt Ponds dominate late-summer melt pond area and exhibit distinct morphological characteristics. DEM-assisted analysis reveals significant stage-dependent differences in local relative elevation. The proposed framework links deep-learning-based stage classification with melt pond morphology and microtopographic associations.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

1 Introduction

Melt ponds on the Arctic sea-ice surface in summer are a key surface feature regulating surface albedo and sea-ice mass balance. Their formation, expansion, drainage, and refreezing are not isolated events, but successive stages in sea-ice surface evolution that influence heat absorption, meltwater redistribution, and the fate of freshwater at the ice–ocean interface (Webster et al., 2022). As the extent of Arctic sea ice has been on the decline for decades, an open Arctic Ocean free of sea ice is projected to occur in the near future (Stroeve and Notz, 2018; Heuzé and Jahn, 2024). In this context, refined observations of melt ponds on summer sea ice are still urgently needed, for the understanding of physical processes at summer sea-ice surface and accurate parameterization in numerical modeling (Lu et al., 2018; Perovich et al., 2008).

Previous studies have primarily focused on melt pond fraction retrieval, geometric characterization, and the relationship between melt ponds and sea-ice microtopography (Aparício et al., 2023; Huang et al., 2016). However, most of these studies still treat melt ponds as a static target, while potential changes in pond surface conditions during melt pond evolution, such as freezing, were neglected (Lee et al., 2020; Aparício et al., 2023). As a result, it remains difficult to establish quantitative links between floe-scale microphysical processes and the regional/basin-scale ablation response of sea ice (Polashenski et al., 2012; Hohenegger et al., 2012; Ding et al., 2020; Niehaus et al., 2025).

Recent advances in low-altitude UAV remote sensing and deep learning have created new opportunities for centimeter-scale melt pond observations. Airborne and UAV imagery have already improved melt pond extraction, sea-ice surface classification, and three-dimensional reconstruction (Miao et al., 2015; Sudakow et al., 2022; Reil et al., 2024; Wang et al., 2018; Zhou et al., 2025), but these developments have mainly enhanced the detection of melt ponds as surface features rather than the interpretation of their evolving surface states (Aparício et al., 2023).

To address these gaps, this study develops a stage-aware framework for Arctic melt pond analysis using high-resolution UAV imagery. Three observable melt pond surface states—Open, Transitional, and Frozen Melt Ponds—are distinguished using the proposed MP-Unet semantic segmentation model. The resulting stage-resolved maps are then used to examine site-level composition, object-level area and shape characteristics, and associations with ridge-related microtopography. Specifically, this study aims

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

to: (1) evaluate the capability of MP-Unet to distinguish melt pond surface states under complex sea-ice background conditions; (2) quantify stage-dependent differences in pond coverage, area distributions, and circularity; and (3) assess whether melt pond stages occupy different local microtopographic positions relative to ridge-like features. By linking stage-aware classification with morphological and topographic analyses, the study provides a process-relevant observational framework for characterizing heterogeneous melt pond conditions during the late Arctic summer.

2 Research Data

2.1 Study Area

According to the NOAA Arctic Report Card (Moon et al., 2024), in the 2024 summer, sea surface temperatures in most Arctic marginal seas were 1–4 °C above the 1991–2020 climatological mean, accompanied by rapid sea-ice retreat. By contrast, the study area in the central Arctic Ocean (85.5°–87.5°N, 142.0°–155.0°E, Fig. 1), remained characterized by thicker multiyear ice floes and relatively high sea-ice concentration of about 90%.

During the first phase of the 14th Chinese National Arctic Research Expedition, R/V Xuelong 2 entered the Arctic Ocean through the Bering Strait on 17 July and reached 80° N on 4 August before sailing northeastward into the central Arctic. Several temporary ice stations were subsequently established on large ice floes for multidisciplinary investigations. In this region, surface melt progressed relatively slowly with the gradual development of melt ponds driven by synoptic processes.

According to meteorological records from the automatic weather station (AWS) aboard R/V Xuelong 2 during 7–15 August, near-surface air temperature was predominantly below freezing, averaging −1.54 °C and ranging from −4.1 to 1.1 °C, while relative humidity remained high (74%–99%). Visibility was generally moderate to good but occasionally fell below 1 km. Station pressure rose to approximately 1012 hPa before decreasing to about 986 hPa on 14 August, indicating the influence of a transient low-pressure system. During the UAV surveys, estimated air temperature ranged from −1.67 to −0.43 °C and wind speed from 2.74 to 10.84 m s^{-1}. The S3 and L1 flights were conducted under particularly strong winds of approximately 10–11 m s^{-1}, and the wind direction shifted markedly among the survey periods. These predominantly cold, humid, and dynamically varying meteorological conditions provide essential

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

context for interpreting melt-pond surface states, freezing processes, and inter-flight differences in the UAV observations.

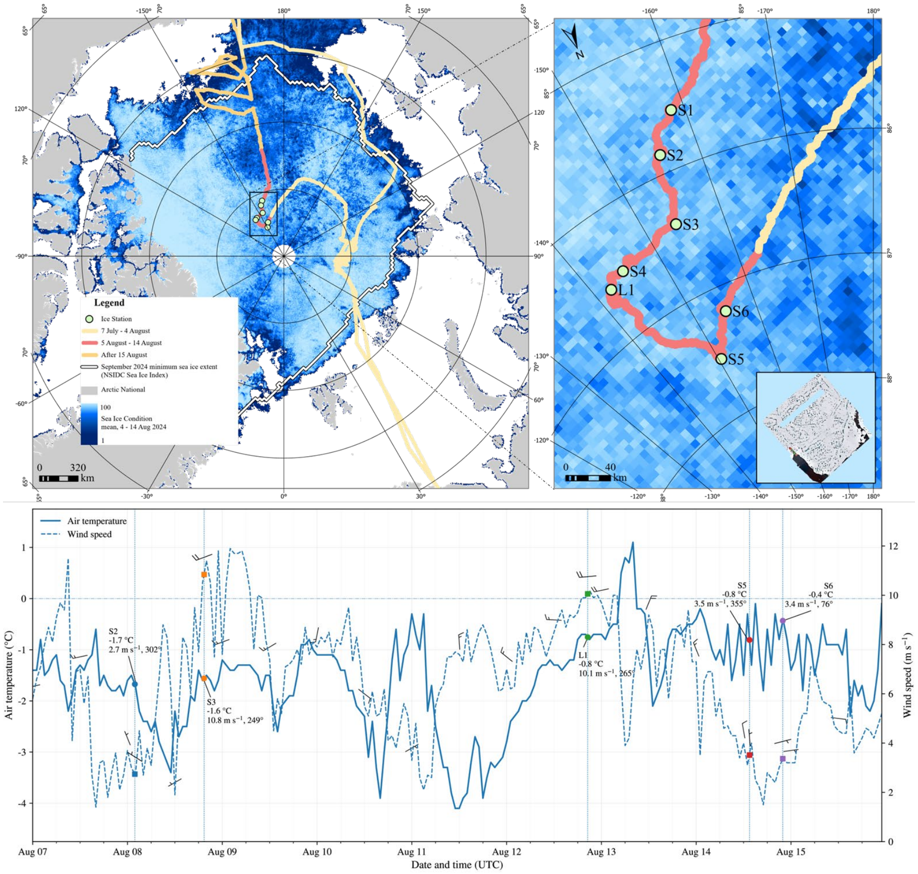


Fig 1. Study area, UAV orthomosaic at Station S3, and AWS meteorological conditions during UAV operations.

2.2 Data

UAV observations were conducted at 5 temporary ice stations on 8, 12, and 14 August 2024, corresponding to the late melt season and a critical period approaching the annual minimum sea-ice extent. A DJI Mavic 3 Classic (M3C) quadrotor was used to acquire optical imagery of the sea-ice surface along parallel flight tracks, with a mean flight speed of approximately 10 m s^{-1}. The M3C platform was equipped with a Hasselblad L2D-20c camera featuring a 4/3-inch CMOS sensor with 20 megapixels, a focal length of 12 mm, and an aperture range of f/2.8–f/11. To ensure image quality over highly reflective snow and ice surfaces, all aerial surveys were conducted using unified camera settings, including an

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

aperture of f/2.8, a shutter speed of 1/1600 s, and an ISO of 100, to obtain imagery with a high signal-to-noise ratio. Flight routes and altitudes were adjusted in real time according to local ice and meteorological conditions to balance spatial coverage and ground-resolution.

A total of seven flights were implemented, covering an effective observation area of approximately 3.7 km² in total (Table 1). High-resolution aerial imagery and video data covering the main floe where the ice-station was established, were acquired. Quality control and basic geometric correction were carried out before further analysis.

Table 1. UAV remote sensing during the CHINARE-14 expedition

Ice Station	Acquisition Time (UTC)	Image size (pixels)	Coverage (km²)	Number of Images
S2	2024-08-08T01:49Z	5280*3956	0.7 km^2	198
S3	2024-08-08T19:19Z	5280*3956	0.8 km^2	337
L1	2024-08-12T20:28Z	3840*2160	0.9 km^2	549
S5	2024-08-14T13:26Z	5280*3956	0.7 km^2	266
S6	2024-08-14T21:49Z	5280*3956	0.6 km^2	228

To synchronously record platform motion, the differential GNSS receiver installed on Xuelong 2 continuously logged ship position (latitude and longitude), heading, and Coordinated Universal Time (UTC) at a sampling frequency of 1 Hz. These data were used to characterize platform motion within the sea-ice reference frame and to constrain subsequent drift correction of the UAV imagery, thereby improving the spatial consistency of the image sequence.

2.3 Drift correction and photogrammetric processing

In the Arctic sea-ice environment, airborne imagery is often affected by systematic spatial offsets caused by sea-ice drift, which appear as time-varying misalignments between consecutive images. These offsets are primarily associated with the advection of drifting sea ice, although minor rotational components may also be present. To reduce their influence on image registration and subsequent melt pond analysis, this study adopted a translation-based drift-correction method using the drift trajectory of the sea-ice floe. Because the research vessel was moored to the sea-ice surface during the field investigation, changes in its GPS position were used as an approximate proxy for the motion of the underlying sea-ice floe (Lund et al., 2018; Krumpen et al., 2021). On this basis, the sea ice within the observed area was approximated as a rigid body over short time scales, and sea ice motion was dominated by translational movement.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

Specifically, the acquisition time of the first image, t_0, was taken as the reference time. For each subsequent image acquired at time t_i, the corresponding ship position was estimated from the shipborne GPS trajectory using cubic-spline interpolation. The relative drift displacement between t_i and t_0 was then represented by the translation vector:

$$\Delta P_i = P_{ship}(t_i) - P_{ship}(t_0) \tag{1}$$

Where $P_{ship}(t)$ denotes the interpolated ship position at time t, and ΔP_i represents the estimated drift displacement of the sea-ice floe relative to the reference time.

The geolocation of each image was subsequently corrected by applying the inverse of the estimated drift vector:

$$P_{corrected} = P_{raw} - \Delta P_i \tag{2}$$

Where P_{raw} and $P_{corrected}$ denote the original and drift-corrected image positions, respectively. This procedure was applied as a whole-image translation to compensate for the bulk translational motion of the sea ice.

Following drift correction, the corrected UAV imagery was processed in Agisoft Metashape Professional (Agisoft LLC, Russia) using a standard Structure-from-Motion (SfM) photogrammetric workflow. The processing pipeline included feature detection and matching, image alignment, bundle adjustment, dense point-cloud reconstruction, mesh generation, and DEM interpolation. During image alignment, the drift-corrected image positions were used as initial camera locations to improve geometric consistency within the common sea-ice reference frame.

The reconstructed dense point cloud was subsequently interpolated to generate a digital elevation model (DEM) representing the local surface topography of the sea-ice floe. Orthorectification and image mosaicking were then performed to produce the final orthomosaic used for melt pond mapping. Consequently, both the orthomosaic and DEM were generated from the same photogrammetric reconstruction workflow and shared a consistent spatial reference.

Because no ground control points were available during the expedition and photogrammetric reconstruction over snow- and ice-covered surfaces is subject to uncertainties associated with low-texture regions, the DEM was used exclusively for analyses of local relative elevation rather than absolute surface elevation. Accordingly, all subsequent topographic analyses focused on elevation differences

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

between individual melt ponds and their surrounding reference areas.

3 Methods

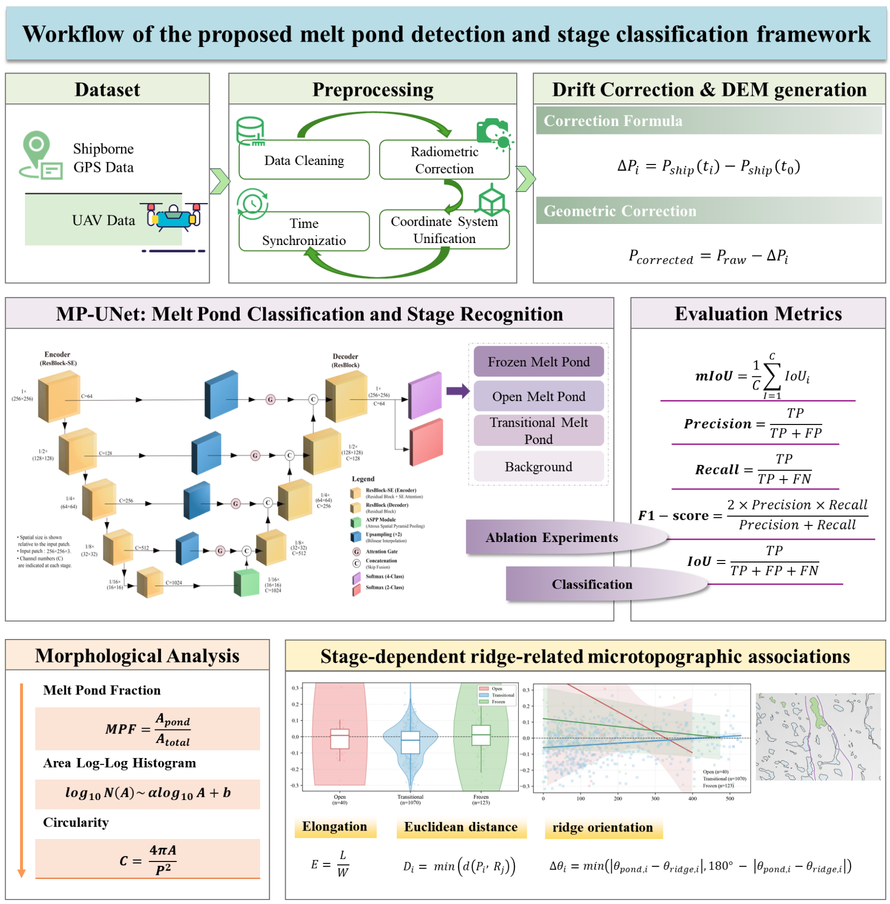


Fig 2. Workflow of the study

Figure 2 illustrates the overall methodological framework of this study for Arctic melt pond identification and stage classification based on UAV observations and deep learning. Using shipborne GPS data and high-resolution UAV imagery, the workflow consists of data preprocessing, sea-ice-drift correction, melt pond stage classification, model performance evaluation, and subsequent morphological and spatial analyses.

Specifically, the raw observations were first preprocessed through data cleaning, temporal interpolation, coordinate unification, radiometric adjustment, and drift correction to ensure cross-source consistency.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

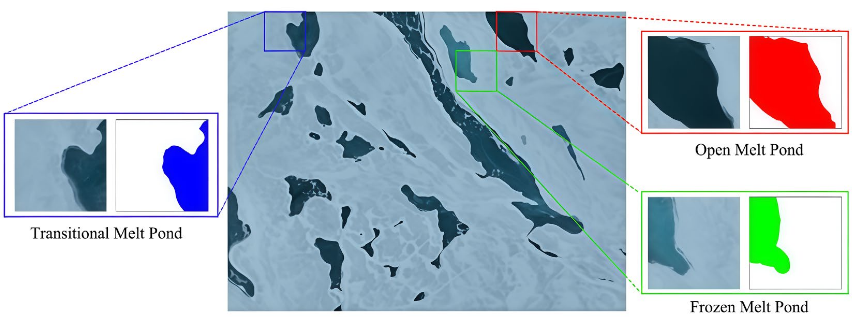
The proposed MP-Unet was then applied to perform pixel-level classification of the sea-ice surface into Frozen Melt Pond, Open Melt Pond, Transitional Melt Pond, and background classes, and its performance was evaluated using Precision, Recall, F1-score, IoU, and mIoU. Based on the classification outputs, further analyses of morphology and spatial organization were conducted to characterize the geometric properties, aggregation patterns, and dominant elongation directions of melt ponds at different stages.

3.1 MP-Unet

3.1.1 Training Dataset

The corrected UAV orthomosaic was manually interpreted and annotated for melt pond classification based on field observation experience and visual interpretation. According to the surface state of melt ponds and their visual appearance in the UAV imagery, the annotated samples were divided into four classes: background, Open Melt Pond, Transitional Melt Pond, and Frozen Melt Pond.

The background class was assigned an attribute code of 0 and an RGB value of (255, 255, 255). It mainly refers to sea-ice surfaces outside melt ponds, including bare ice, snow-covered ice, and other non-pond areas. Open Melt Ponds were assigned an attribute code of 1 and an RGB value of (255, 0, 0). This class mainly corresponds to ponds without visible ice cover on the surface, which typically appear dark blue to deep blue in the UAV imagery. Transitional Melt Ponds were assigned an attribute code of 2 and an RGB value of (0, 0, 255). This class represents a transitional state between Open Melt Ponds and Frozen Melt Ponds, often characterized by mixed blue tones, partial surface ice cover, or heterogeneous surface textures, indicating the unstable and transitional nature of the pond surface. Frozen Melt Ponds were assigned an attribute code of 3 and an RGB value of (0, 255, 0). This class mainly refers to melt ponds whose surfaces are covered by continuous or nearly continuous ice.



https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

Fig 3. Annotation examples and representative subset image

A set of RGB-derived optical features was extracted from the manually annotated images (Fig. 4). Besides the DNs of the R, G, and B channels, brightness, $L*$, and saturation were used to describe class separability. Brightness represents the apparent light–dark intensity of the RGB image, $L*$ denotes perceptual lightness in the CIE 1976 $L*a*b*$ colour space, and saturation describes the departure of a colour from an achromatic grey or white appearance in HSV/HSB-type colour models (CIE, 2007; Smith, A. R. 1978). The probability density distributions of the R, G, and B channels (Fig. 4a–c), together with the boxplots of brightness, $L*$, and saturation (Fig. 4d–f), indicate that the three manually defined melt pond classes are visually distinguishable and separable in optical-feature space.

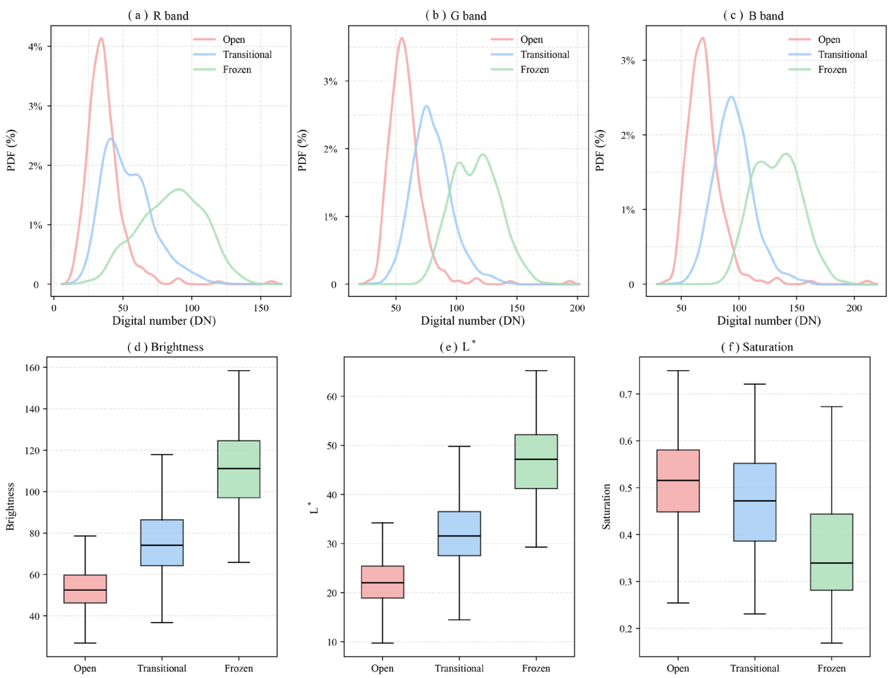


Fig 4. Object-level optical-feature distributions of the three melt pond classes. (a – c) Probability density distributions of DN in the R, G, and B channels. (d – f) Boxplots of brightness, $L*$, and saturation.

The constructed dataset contains a total of 70 annotated images, of which 65 were used for training and 5 for testing and qualitative visual evaluation. The dataset was divided at the image level rather than at the patch level to prevent spatially adjacent patches from the same source image from appearing in both training and testing sets. Considering GPU memory constraints, the need for stable batch-based training, and the representation of fine-scale textures and boundary structures, image patches of 256×256 pixels

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

were adopted as the basic input unit for both training and inference (Fig. 3). After sliding-window cropping of all annotated images, 18,120 image–label patch pairs were generated for model training.

To improve model generalization and reduce the risk of overfitting, online random data augmentation was applied during training. Unlike offline augmentation, which explicitly duplicates samples, online augmentation dynamically applies random geometric and photometric transformations to the same patch at each iteration, thereby expanding the effective sample distribution without increasing storage burden. The augmentation operations included common semantic-segmentation transformations such as rotation, flipping, scaling, cropping, noise perturbation, and color adjustment. This strategy improves the model's adaptability to images with variations in orientation, scale, and local illumination, and thus enhances robustness of trained model under complex sea-ice surface conditions.

To ensure comparability among subsequent ablation experiments, all experiments used the same data split, and the random state was explicitly fixed by using the same random-seed in the DataLoader workers. Meanwhile, all experimental groups shared the same random-sequence settings to minimize performance fluctuations caused by stochastic perturbations and to improve the interpretability of performance differences introduced by different model components.

3.1.2 Network Architecture Design

Built upon the classical U-Net encoder–decoder architecture (Ronneberger et al., 2015), a four-class semantic segmentation of summer sea-ice surfaces, namely MP-Unet, was designed to address the challenges posed by the similarity between melt ponds and background textures, ambiguous class boundaries, and pronounced variations in melt pond scale on UAV imagery. The network retains the multi-scale feature fusion advantage of the U-shaped architecture while introducing several task-oriented modifications to key modules (Fig. 5). Specifically, MP-Unet incorporates residual blocks-SE with channel attention in the encoder, an atrous spatial pyramid pooling module in the bottleneck, an attention gating mechanism in the skip connections, and an auxiliary binary classification head at the end of the decoder, thereby enhancing the representation and segmentation of melt ponds at different stages under complex sea-ice background conditions.

First, Residual learning (He et al., 2016) was adopted by replacing the Double Conv module in the original U-Net with a residual convolutional block to capture subtle textures among different melt pond stages, weak-contrast boundaries, and local variations in surface state. In addition, a Squeeze-and-

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

Excitation (SE) channel attention mechanism (Hu et al., 2020) was embedded into the residual block-SE to explicitly model inter-channel dependencies and adaptively recalibrate feature channels, thereby strengthening discriminative information related to melt pond boundaries, local textures, and surface freezing states. Let the input feature be x; then the output of the residual module can be expressed as Equation (3):

$$y = \sigma(F(x) + W_S x) \tag{3}$$

where $F(x)$ denotes the residual mapping learned by convolutional transformations, $W_S x$ is the linear mapping of the shortcut branch, and σ is the nonlinear activation function. On this basis, the SE module further reweights residual features along the channel dimension, improving gradient propagation in deep networks while enhancing the network response to key semantic features.

Second, given the substantial scale variation of melt pond targets, relying solely on local convolutional receptive fields is insufficient to capture their multi-scale contextual information, especially in complex backgrounds and near boundary regions where misclassification is more likely to occur. To address this issue, an Atrous Spatial Pyramid Pooling (ASPP) module (Chen et al., 2018) was introduced into the bottleneck layer. By combining parallel convolutional branches with different atrous rates and a global average pooling branch, the ASPP module jointly models contextual information at multiple scales. This design effectively enlarges the receptive field without substantially increasing the number of parameters, thereby improving the model's ability to perceive melt pond morphology at different scales as well as its surrounding environmental context, and ultimately enhancing multi-class representation in complex scenes.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

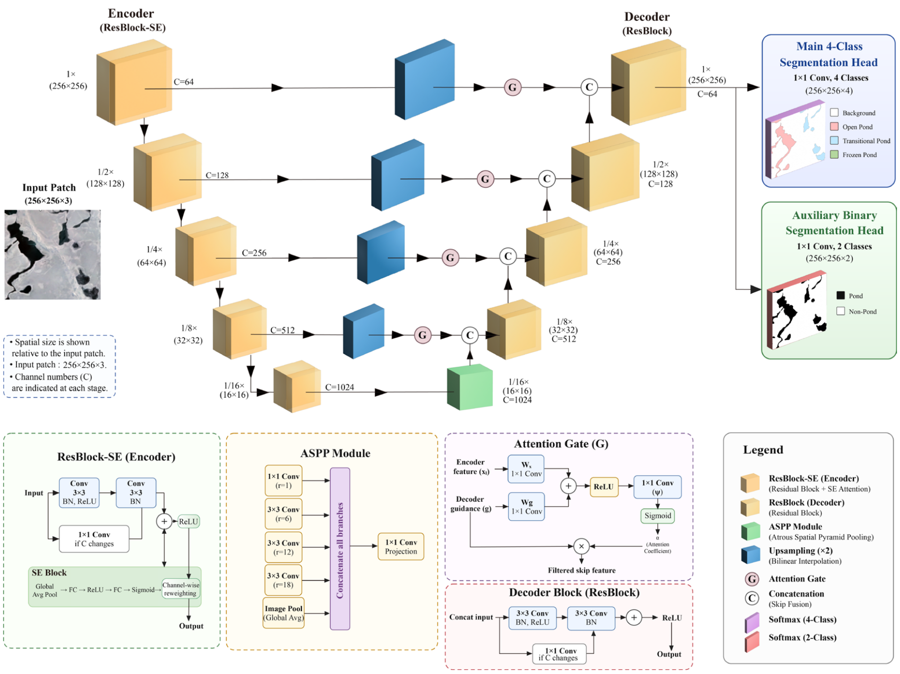


Fig 5. Model Architecture Diagram

Third, in the decoder, the original U-Net directly fuses low-level encoder features with high-level decoder features through skip connections. Although this design helps recover spatial details, it may also pass background textures and noise irrelevant to the target into the decoder, thereby weakening classification discriminability. To alleviate this problem, an Attention Gate module (Oktay et al., 2018) was introduced into the skip connections, allowing high-level semantic information from the decoder to selectively filter low-level encoder features. This suppresses irrelevant background responses while enhancing effective features related to melt ponds. The gated skip features are then fused with the upsampled features and further refined by residual convolutional modules, thereby improving boundary recovery and class discrimination, particularly for weak-boundary targets such as Transitional Melt Ponds.

In addition, to further reduce false positives in the background and improve the overall separability of foreground regions, an auxiliary binary classification head was introduced at the end of the decoder. This branch divides the sea-ice surface into two classes, namely melt pond and non-melt pond, and is jointly supervised with the main four-class branch during training. Rather than replacing the main task, this auxiliary branch serves as an additional constraint, guiding the network to maintain overall sensitivity to foreground regions while learning fine-grained inter-class differences, thereby reducing mis-

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

classification under complex sea-ice background conditions. Ultimately, the main branch outputs pixel-level classification results for background, Open Melt Ponds, Transitional Melt Ponds, and Frozen Melt Ponds, whereas the auxiliary branch improves the stability and robustness of foreground recognition.

Overall, MP-Unet is a task-oriented architecture specifically designed for Arctic melt pond stage classification through coordinated enhancement of feature extraction, multi-scale contextual modeling, skip-feature selection, and background suppression. Through the combined effects of ResBlock-SE, ASPP, Attention Gate, and the auxiliary binary classification head, the model can more effectively cope with the complex background, subtle inter-class differences, and potential scale variation of Arctic sea-ice surfaces, thereby improving the accuracy and stability of multi-stage melt pond classification.

3.1.3 Loss Function and Training Strategy

For loss design, the main branch was optimized using a weighted combination of weighted cross-entropy loss and Dice loss. The weighted cross-entropy term enhances pixel-level class discrimination and alleviates the effects of class imbalance through category-specific weights, whereas the Dice loss improves the learning of small targets and boundary regions. The loss function for the main branch is defined in Equation (4).

$$L_{main} = 0.7L_{CE} + 0.3L_{Dice} \tag{4}$$

L_{CE} denotes the multi-class cross-entropy loss with class weights, and L_{Dice} denotes the multi-class Dice loss. The class weights were set to [1.6,2.0,2.0,2.0] for background, Open Melt Ponds, Transitional Melt Ponds, and Frozen Melt Ponds, respectively, to place greater emphasis on foreground classes.

The auxiliary branch was supervised using a binary cross-entropy loss for the melt pond/non-melt pond classification task, and its loss is defined as Equation (5):

$$L_{aux} = L_{CE}^{binary} \tag{5}$$

Accordingly, the total loss function of the model can be written as Equation (6):

$$L = L_{main} + 0.3L_{aux} = 0.7L_{CE} + 0.3L_{Dice} + 0.3L_{aux} \tag{6}$$

By jointly optimizing the main and auxiliary tasks, the network can maintain stable sensitivity to the overall melt pond foreground while learning fine-grained stage differences, thereby improving classification accuracy and robustness under complex sea-ice background conditions.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

All experiments were carried out on a workstation equipped with an NVIDIA RTX 4090 GPU (24 GB). The software environment consisted of CUDA 12.1, Python 3.12, and PyTorch 2.3.0. Model training was performed using the AdamW optimizer, which incorporates decoupled weight decay into adaptive learning-rate optimization and thus improves both training stability and model generalization. Based on preliminary tests and hardware constraints, the main training hyperparameters used in this study are summarized in Table 2.

Table 2. Hyperparameter Configuration

hyperparameters	Value
Maximum number of training epochs	1000
Batch size	16
Learning Rate	0.0001
Weight Decay	0.0005
Early Stopping	10

Input samples were uniformly cropped into 256 × 256 patches, and the base number of network channels was set to 32. Mixed-precision training was adopted to improve training efficiency and GPU memory utilization. To reduce overfitting and enhance training stability, online random data augmentation and an early stopping strategy were jointly employed, with the patience parameter set to 10.

To improve the model's adaptability to complex dark-background regions, a dark-background-aware sampling strategy was introduced during training. Specifically, the proportion of dark background in each sample was first estimated based on grayscale statistics, and the training samples were then stratified accordingly. In each batch, dark-background patches were included at a fixed proportion to improve the model's ability to distinguish low-intensity background regions from melt-pond-like areas. Without altering the overall data split, this strategy increased the coverage of complex background conditions in the training data, thereby enhancing model robustness in real sea-ice scenes.

Overall, the adopted joint loss design, standardized training configuration, and background-aware sampling and model-selection strategies improved the model's adaptability to class imbalance, weak boundaries, and complex background conditions, while also ensuring fairness, reproducibility, and interpretability of the experimental results. These settings provide a reliable basis for subsequent model evaluation and result analysis.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

3.1.4 Evaluation Metrics

To quantitatively evaluate the classification performance of the model for melt ponds at different stages, Precision, Recall, F1-score, Intersection over Union (IoU), and mean Intersection over Union (mIoU) were adopted as the primary evaluation metrics. Among them, Precision, Recall, F1-score, and IoU were mainly used to assess the model's ability to identify target classes, while mIoU was used to characterize the overall performance of the model in the multi-class segmentation task.

For a given class, let TP (True Positive) denote the number of pixels correctly predicted as that class, FP (False Positive) denote the number of pixels incorrectly predicted as that class, and FN (False Negative) denote the number of pixels belonging to that class but not correctly identified. The evaluation metrics are defined as follows.

Precision represents the proportion of pixels predicted as the target class that belongs to that class. It is used to measure the accuracy of positive predictions, as shown in Equation (7).

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

Recall represents the proportion of pixels belonging to the target class that are correctly identified by the model. It is used to measure the detection capability of the model for the target class, as shown in Equation (8).

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

The F1-score is the harmonic mean of Precision and Recall and is used to comprehensively evaluate the balance between precision and completeness, as shown in Equation (9).

$$F1 - \text{score} = \frac{2 \times Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FN + FP} \tag{9}$$

The IoU represents the degree of overlap between the predicted region and the ground-truth region, and is one of the commonly most used metrics in semantic segmentation, as shown in Equation (10).

$$IoU = \frac{TP}{TP + FP + FN} \tag{10}$$

mIoU is defined as the average IoU over all classes and is used to evaluate the overall performance of the model in the multi-class segmentation task. Let C denote the total number of classes, and let IoU denote the IoU of the i-th class, as shown in Equation (11).

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

$$mIoU = \frac{1}{C}\sum_{I=1}^{C} IoU_i \tag{11}$$

3.2 Melt Pond parameters

After the classification, further quantitative analyses were conducted from two complementary perspectives: individual geometric characteristics and collective spatial organization. The former was used to characterize melt pond area, boundary complexity, and shape properties, whereas the latter was used to reveal local-scale aggregation patterns, dominant elongation directions, and directional consistency. By combining morphological and spatial analyses, the geometric attributes and spatial organization of melt ponds at different stages can be characterized more systematically, thereby providing quantitative support for understanding melt pond evolution and its environmental response.

(1) Melt Pond Fraction (MPF)

Melt Pond Fraction (MPF) describes the proportion of melt pond area relative to the total area within a given analysis region and is a fundamental metric for characterizing melt pond coverage. MPF can be calculated separately for different melt pond stages to quantify the relative proportions of Frozen Melt Ponds, Open Melt Ponds, and Transitional Melt Ponds in the study area, as shown in Equation (12):

$$MPF = \frac{A_{pond}}{A_{total}} \tag{12}$$

where A_{pond} is the area of melt ponds belonging to the target class, and A_{total} is the total area of the analysis region. A larger MPF indicates greater areal coverage of that melt pond class within the region.

(2) Area distribution

To characterize the scale-dependent distribution of melt pond areas, the relationship between the area of an individual melt pond, S, and its corresponding number of ponds, $N(S)$, was analyzed and fitted in logarithmic space. Previous studies have shown that melt pond area distributions often exhibit power-law behavior over a certain range of scales (Huang et al., 2016; Popović et al., 2018), as expressed in Equation (13):

$$log_{10} N(S) = \alpha log_{10} S + b \tag{13}$$

where S denotes the representative area of a logarithmically spaced area bin, $N(S)$ denotes the number of melt ponds within that bin, α is the power-law exponent, and b is a constant term. This metric was used to describe area-distribution patterns across different scales and to examine whether scale-dependent

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

differences exist among melt ponds at different stages.

(3) Circularity

Circularity measures the closeness of a melt pond shape to an ideal circle and is an important indicator of shape regularity, as shown in Equation (14):

$$C = \frac{4\pi A}{P^2} \quad (14)$$

where A is the melt pond area and P is the melt pond perimeter. In theory, $C \in [0,1]$. When C=1, the object is a perfect circle. Smaller values of C indicate that the melt pond is more elongated, fragmented, or irregular in boundary shape.

The perimeter-to-area ratio was used only as an auxiliary metric for interpreting boundary complexity. For melt ponds with similar areas, a larger perimeter-to-area ratio suggests a more convoluted or fragmented boundary. In this study, circularity was therefore treated as the main indicator of shape regularity, whereas the perimeter-to-area ratio was used to provide supplementary evidence for edge complexity.

3.3 Continuous analysis of stage-dependent ridge-related microtopographic associations

(1) Elongation

To examine the associations between ridge-related microtopography and melt pond characteristic, the minimum rotated bounding rectangle was fitted to each melt pond polygon. The elongation ratio was defined as the ratio of the major axis length to the minor axis length:

$$E = \frac{L}{W} \quad (15)$$

where L and W denote the lengths of the major and minor axes of the minimum rotated rectangle, respectively. Larger values of E indicate increasingly elongated pond morphology, whereas values approaching 1 indicate a more equant shape.

(2) Orientation

The dominant orientation of each melt pond was determined from the major axis of its minimum rotated rectangle. Orientation was expressed as the clockwise angle from geographic north, ranging from 0° to 180°. This orientation characterizes the principal elongation direction of each melt pond and was subsequently compared with the orientation of the nearest ridge-like feature.

(3) DEM-based ridge–pond association analysis

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

To investigate the influence of ridge-related microtopography on melt pond organization, DEM data and manually digitized ridge-like linear features were incorporated as auxiliary topographic information. All melt pond polygons, ridge-like features, and DEM data were transformed into the same polar stereographic coordinate system (EPSG:3413). Ridge-like features were interpreted from the DEM hillshade and UAV orthomosaic and were regarded as indicators of local microtopography rather than exact ice-ridge boundaries.

For each melt pond, the Euclidean distance from the pond boundary to the nearest ridge-like feature was calculated as

$$D_i = min(d(P_i, R_j)) \tag{16}$$

where D_i denotes the shortest distance between melt pond P_i and all ridge-like features R_j.

Local topographic position was quantified using DEM-derived relative elevation. For each melt pond, the median elevation inside the pond was compared with the median elevation of a surrounding annular zone extending from 2 to 5 m outside the pond boundary after excluding all neighboring pond areas:

$$\Delta h_i = h_{pond,i} - h_{ring,i} \tag{17}$$

where $h_{pond,i}$ represents the median elevation within pond *i*, and $h_{ring,i}$ represents the median elevation of the surrounding reference ring. Negative values indicate that the pond occupies a locally depressed topographic position.

To evaluate directional coupling between melt ponds and ridge-like features, only elongated ponds (elongation ratio $\geqslant$ 1.5) were included. The local orientation of the nearest ridge was estimated from a short line segment centered on the nearest point along the ridge. The angular difference between pond orientation and ridge orientation was calculated as

$$\Delta\theta_i = min(|\theta_{pond,i} - \theta_{ridge,i}|, 180° - |\theta_{pond,i} - \theta_{ridge,i}|) \tag{18}$$

where $\Delta\theta_i$ ranges from 0° to 90°, with smaller values indicating stronger parallel alignment between the pond major axis and the local ridge orientation.

Finally, statistical analyses were conducted to evaluate stage-dependent associations between melt ponds and ridge-related microtopography. Ordinary least squares (OLS) regression models were first used to examine the relationships between ridge distance and (i) melt pond area and (ii) local relative elevation

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

while accounting for melt pond stage and the interaction between ridge distance and stage. This interaction term was introduced to assess whether ridge-related topographic associations differed among Open, Transitional, and Frozen Melt Ponds.

To further quantify stage-dependent differences in microtopographic position, the distributions of ridge distance and local relative elevation were compared among the three melt pond stages using the non-parametric Kruskal–Wallis test. When significant overall differences were detected, pairwise Mann–Whitney U tests with Holm correction were performed for multiple comparisons.

Directional coupling between elongated melt ponds and nearby ridge-like features was evaluated using OLS regression and Spearman's rank correlation analysis. Only melt ponds with an elongation ratio greater than or equal to 1.5 were included in the directional analysis.

4 Results

4.1 Model Performance Evaluation

Ablation experiments were conducted to evaluate the contributions of the Residual Block-SE, ASPP module, Attention Gate, and the complete MP-Unet to melt pond segmentation performance. The baseline U-Net, three single-module variants, and the complete MP-Unet were evaluated using the same test set described above. The quantitative results are summarized in Table 3.

Table 3. Quantitative performance comparison of the baseline, single-module variants, and the complete MP-Unet.

Indicator \ Model	Baseline	＋Residual Block-SE	＋ASPP	＋Attention Gate	MP-Unet
Validation loss	0.1780	0.1754	0.2531	0.1733	**0.1703**
mIoU	0.7288	0.7361	0.6487	0.7430	**0.7466**
F1-Score	0.9387	0.9396	0.9126	0.9393	**0.9440**
Precision	0.9300	0.9219	0.9161	0.9274	**0.9365**
Recall	0.9475	**0.9580**	0.9090	0.9515	0.9515
Background false positive rate	0.0073	0.0083	0.0085	0.0076	**0.0066**
IoU_Open	0.5256	0.5600	0.4050	0.5617	**0.5897**
IoU_Trans	0.7479	0.7491	0.6847	**0.7587**	0.7544

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

IoU_Frozen	0.6543	0.6476	0.5232	**0.6639**	0.6539
IoU_Bg	0.9875	0.9875	0.9824	0.9875	**0.9886**

As shown in Table 3, compared with the baseline U-Net, introducing the Residual Block-SE improved the overall segmentation performance, with the mIoU increasing from 0.7288 to 0.7361 and the F1-score increasing from 0.9387 to 0.9396. The Recall also increased to 0.9580, indicating that residual learning with channel attention enhances feature representation and improves the sensitivity of the network to melt pond regions. Although the Precision and the IoU of the Frozen Melt Pond class decreased slightly, the overall segmentation performance was consistently improved.

The three single-module variants exhibited different performance characteristics.Among the three single-module variants, the model incorporating the Attention Gate achieved the best overall performance. The mIoU increased to 0.7430, while the IoU of the Transitional Melt Pond class reached 0.7587 and the IoU of the Frozen Melt Pond class increased to 0.6639, both representing the highest values among the single-module configurations. These results indicate that attention-guided skip connections effectively suppress irrelevant background responses while preserving discriminative spatial features, thereby improving the segmentation of melt ponds with weak boundaries and heterogeneous surface textures.

In contrast, the model incorporating the ASPP module alone showed noticeably weaker performance, with the mIoU decreasing to 0.6487 and the IoU values of Open, Transitional, and Frozen Melt Ponds dropping to 0.4050, 0.6847, and 0.5232, respectively. This suggests that multi-scale contextual aggregation alone is insufficient to alleviate class confusion under complex sea-ice background conditions. Without complementary mechanisms for feature refinement and attention-based selection, the enlarged receptive field may introduce redundant contextual information, resulting in degraded discrimination between visually similar surface classes.

The complete MP-Unet, which integrates the Residual Block-SE, ASPP module, Attention Gate, and auxiliary binary segmentation head, achieved the best overall performance among all configurations. Specifically, it produced the lowest validation loss (0.1703), the highest mIoU (0.7466), the highest F1-score (0.9440), the highest Precision (0.9365), and the lowest background false positive rate (0.0066). In addition, the complete model achieved the highest IoU for the Open Melt Pond class (0.5897) and the Background class (0.9886), demonstrating an enhanced capability to distinguish subtle texture differences and weak boundary features under complex Arctic sea-ice conditions. Although the IoU values of the Transitional and Frozen Melt Pond classes were slightly lower than those achieved by the

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

Attention Gate-only variant, the complete MP-Unet provided the most balanced segmentation performance across all categories.

Overall, the ablation study demonstrates that the proposed modules play complementary rather than independent roles in improving segmentation performance. The Residual Block-SE primarily enhances low-level feature representation, the Attention Gate improves the selective fusion of encoder and decoder features, and the ASPP module provides richer multi-scale contextual information. Although the ASPP module alone did not improve segmentation accuracy, its integration with the other proposed modules contributed to the best overall performance of the complete MP-Unet. These findings indicate that robust stage-aware melt pond segmentation is achieved through the collaborative effects of complementary architectural components rather than by any single module alone.

4.2 Classification Results

Figure 6 presents the four-class melt pond classification results in several representative sea-ice scenes, illustrating the model's ability to identify Frozen Melt Ponds, Open Melt Ponds, Transitional Melt Ponds, and background. Overall, MP-Unet achieved good classification accuracy and spatial consistency across different scenarios and was able to adapt to the morphological and textural variability of melt ponds at different stages under complex sea-ice background conditions.

In the scene dominated by Frozen Melt Ponds (S6), the model successfully identified fragmented and spatially dispersed frozen melt ponds. The classification results correspond well to the light-blue frozen features visible in the original imagery, and the boundaries were delineated clearly. In the scene dominated by Open Melt Ponds (S2), the model effectively distinguished large connected open melt ponds from the surrounding transitional melt ponds, producing segmentation results with good spatial continuity and relatively accurate boundary localization. For melt ponds containing local internal heterogeneity but dominated by one overall stage, the model also showed stable category discrimination. In the scene dominated by Transitional Melt Ponds (S3), the model was able to identify transitional melt ponds with complex morphology and diverse scales, while maintaining good classification consistency in regions with thin ice cover or locally frozen surface features. This indicates a strong discriminative ability for melt ponds in intermediate stages. In the more complex scene containing inter-ice channels and open-water areas (S5), the model effectively suppressed interference from non-target water bodies and accurately identified small melt ponds located along channel margins, without showing obvious class

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

confusion.

The local enlarged views further demonstrate that the model performed stably in melt pond boundary delineation and class separation. Boundaries between different categories were generally clear, with no obvious over-segmentation or under-segmentation. Overall, MP-Unet was able to accurately identify melt ponds at different stages across diverse sea-ice scenes, providing a reliable pixel-level segmentation basis for subsequent morphological and spatial analyses.

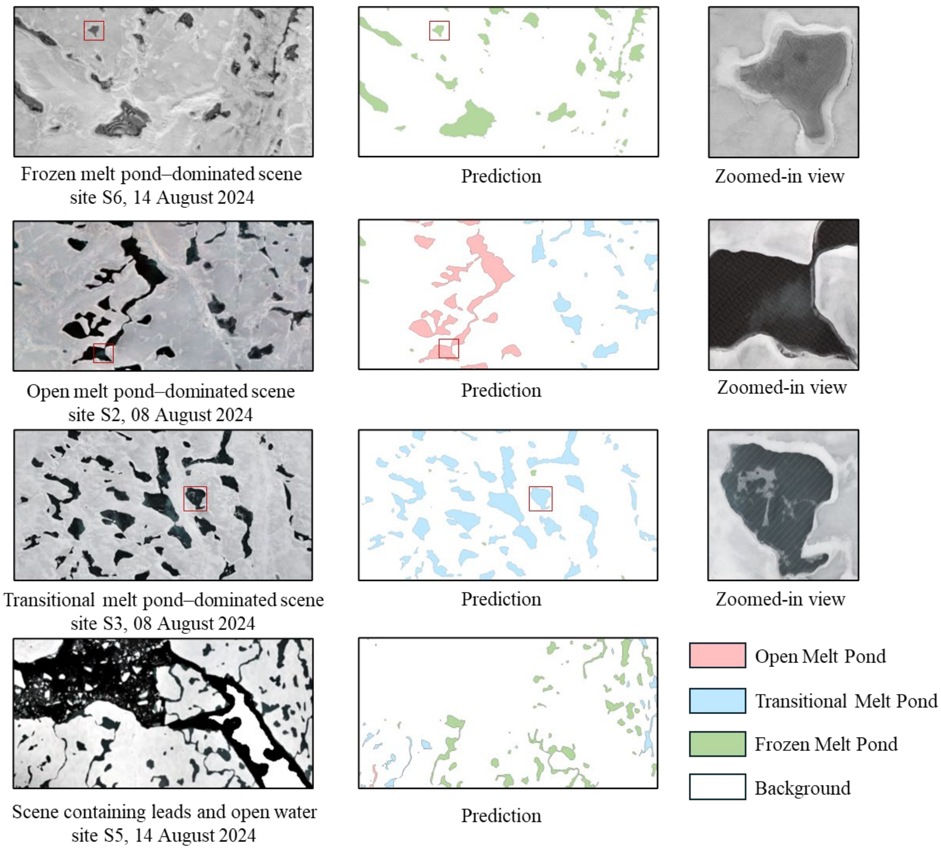


Fig. 6. Melt Pond Classification Results

4.3 Stage composition and object-level morphological differentiation

A multidimensional analysis was conducted to characterize melt pond stage composition, area–frequency scaling, and circularity across five sites and the pooled sample (Fig. 7).

Melt pond fraction and stage composition exhibit pronounced spatial heterogeneity among the five sites. S5 has the highest melt pond fraction, reaching approximately 15.4%, followed by S3 at 9.4% and S2 at

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

4.7%. Considerably lower values occur at L1 and S6, where the melt pond fractions are approximately 2.6% and 2.3%, respectively. Transitional Melt Ponds dominate the total pond area at S2, S3, S5, and L1, accounting for approximately 62.0%, 89.4%, 90.0%, and 54.9%, respectively. In the pooled sample, Transitional Melt Ponds account for about 80.3% of the total pond area, indicating that transitional surface states are the dominant regional condition. Open Melt Ponds make relatively large contributions at S2 and L1, accounting for approximately 26.9% and 30.6%, whereas their contributions are very limited at S3 and S5.

The widespread dominance of Transitional Melt Ponds is consistent with the highly variable surface-energy conditions during the late Arctic summer. Shipboard automatic weather station observations from 5 to 15 August indicated persistently subfreezing but near-melting air temperatures of approximately −2.15 to −0.93 °C, high relative humidity of 88.85%–95.97%, variable wind forcing, and a decrease in atmospheric pressure from approximately 1000 to 986 hPa. Under these conditions, radiative heating may continue to support surface melting, whereas short-term reductions in radiation and enhanced turbulent heat loss may promote partial refreezing. S6 was the only site dominated by Frozen Melt Ponds. This difference is broadly consistent with the regional atmospheric conditions during the survey period but may also reflect local differences in pond properties and microtopography.

The log–log area–frequency curves show a common scale-dependent pattern across the five sites. Small and intermediate ponds are abundant, whereas the number of ponds decreases substantially toward larger area classes. Nevertheless, the distributions extend across several orders of magnitude and retain a sparse large-area tail, indicating the presence of a limited number of exceptionally large ponds. The full-range fitted exponents vary from $\alpha=-0.28$ at S3 to $\alpha=-0.48$ at S6, while the pooled sample yields $\alpha=-0.55$. The shallower slopes at S2 and S3 indicate a slower decrease in pond number with increasing area, whereas the steeper slopes at S6 and L1 indicate a more rapid decline toward the large-area end.

The coefficients of determination range from $R^2=0.42$ to 0.68. The pooled sample has the highest value, with $\alpha=-0.55$ and $R^2=0.68$, followed by S6, L1, and S5. By contrast, S2 and S3 show lower fitting performance, with $R^2=0.47$ and 0.42, respectively. These results suggest that the area – frequency distributions exhibit a general power-law-like decline but cannot be represented by a single scaling relationship with equal strength at all sites. The intermediate-area plateaus and the subsequent rapid decrease toward the large-area tail indicate that local pond coalescence, drainage, fragmentation, and stage conversion may modify the idealized scaling pattern.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

In the pooled sample, Frozen Melt Ponds are shifted toward higher circularity values and are concentrated mainly in the approximately 0.55–0.75 range, indicating comparatively compact and regular shapes. Open Melt Ponds display a broader distribution extending toward lower circularity values, consistent with greater boundary irregularity and geometric complexity. Transitional Melt Ponds generally occupy an intermediate and broader range. This stage-related pattern is evident at S2, S6, and L1, although it is not uniform across all sites. At S5, for example, both Frozen and Open Melt Ponds show relatively high circularity, whereas Transitional Melt Ponds exhibit a broader distribution. The S3 distributions are also partly multimodal. These departures indicate that stage-dependent differences are not consistent across all sites.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

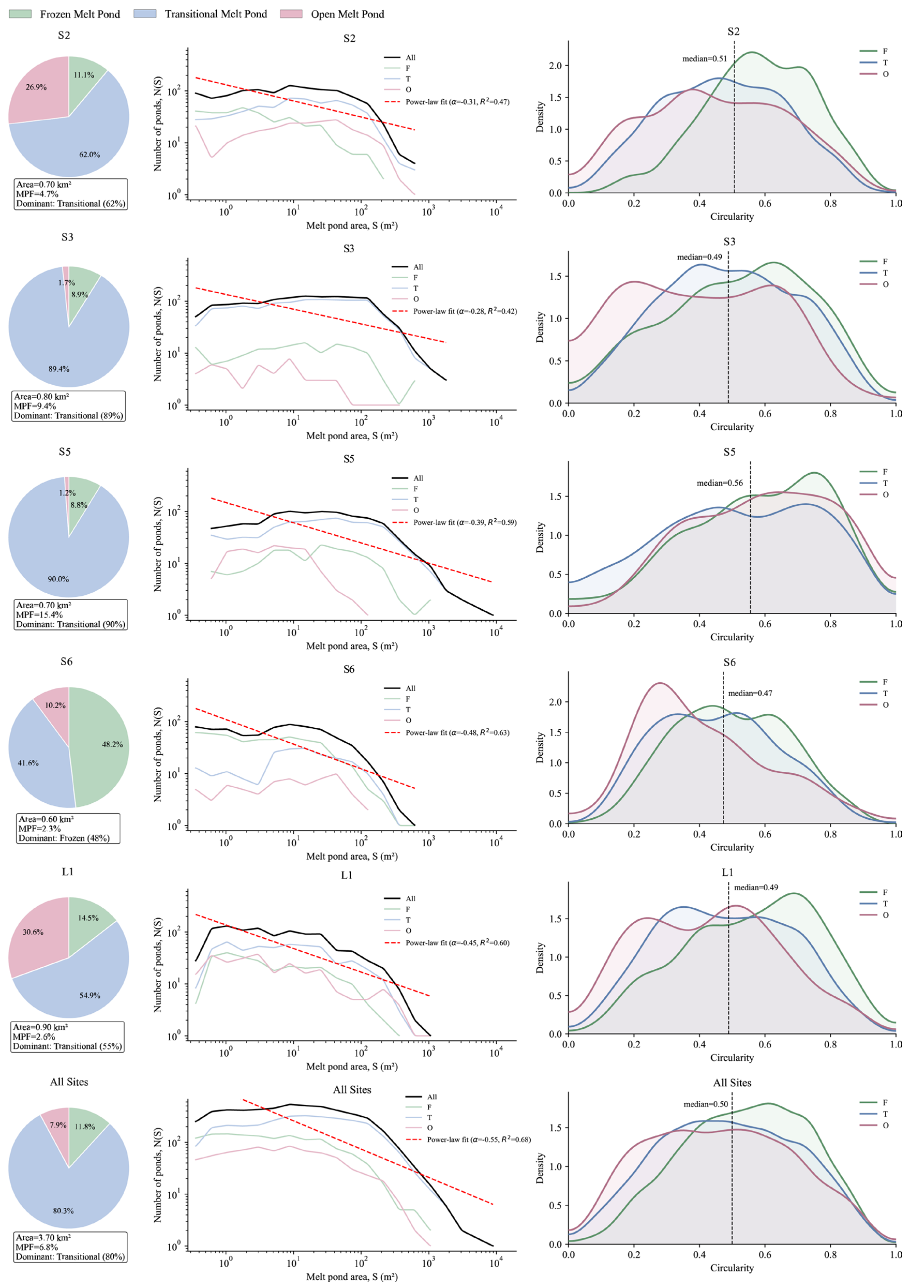


Fig 7. Stage composition, area–frequency scaling, and circularity distributions of melt ponds across sites S2, S3, S5, S6, and L1.

The same-day station pairs further indicate that melt pond composition cannot be explained by observation date or latitude alone. Although S3 is located slightly farther north than S2 and exhibits a

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

stronger dominance of Transitional Melt Ponds, the S5–S6 pair shows the opposite spatial tendency: S5 is farther north than S6 but has a substantially higher melt pond fraction and much stronger Transitional Melt Pond dominance. The observed spatial differences are therefore more likely related to local ice-surface conditions, short-term thermal variability, and microtopographic or ridge-related settings than to a simple latitudinal gradient.

4.4 Stage-dependent ridge-related microtopographic associations

The spatial relationships between melt ponds and ridge-related microtopography are illustrated in Fig. 8. Melt ponds of all developmental stages were widely distributed throughout the study area and generally occurred in the vicinity of ridge-like linear features (Fig. 8a). However, visual inspection alone did not reveal obvious stage-specific clustering with respect to ridge proximity.

To further evaluate stage-dependent microtopographic positioning, local relative elevation was compared among Open, Transitional, and Frozen Melt Ponds (Fig. 8b). Significant differences were detected among the three stages (Kruskal–Wallis test, $p < 0.001$). Transitional Melt Ponds occupied significantly lower local topographic positions than Frozen Melt Ponds, whereas Open Melt Ponds showed intermediate characteristics. This result indicates that melt pond stage is associated with local microtopographic position rather than with large-scale spatial redistribution relative to ridge-like features.

The relationship between local relative elevation and ridge distance further demonstrated stage-dependent differences (Fig. 8c). Transitional Melt Ponds exhibited a weak but significant positive relationship between ridge distance and relative elevation, suggesting that ponds located closer to ridge-like features tend to occupy locally depressed topographic positions. In contrast, Open and Frozen Melt Ponds showed no significant relationship between ridge distance and local relative elevation. Although the interaction between ridge distance and pond stage did not reach conventional statistical significance, the observed stage-specific trends consistently indicate that Transitional Melt Ponds are more sensitive to ridge-related microtopographic variation than the other two stages.

The distributions of ridge proximity among the three melt pond stages are provided in the Fig. A1 in Appendix A. No significant differences in ridge distance were detected among Open, Transitional, and Frozen Melt Ponds, indicating that the observed differences in local relative elevation cannot be attributed simply to variations in ridge proximity. Instead, the results suggest that ridge-related microtopography primarily influences the local topographic setting occupied by melt ponds during the

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

transitional stage rather than directly controlling their spatial occurrence or planform morphology.

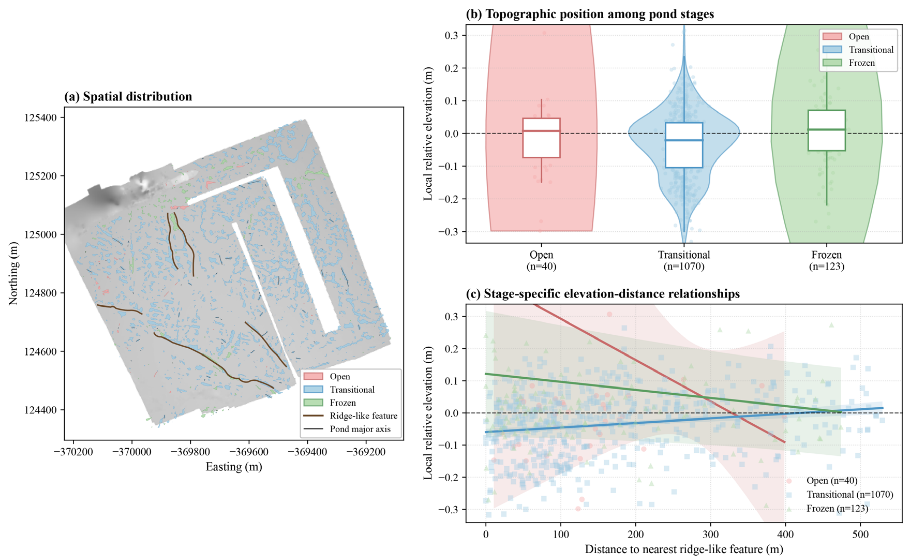


Fig 8. Microtopographic positioning of different melt pond stages relative to ridge-like features.

5 Discussion

The results demonstrate that stage-aware semantic segmentation can extend UAV-based melt pond mapping beyond the conventional distinction between ponded and non-ponded sea-ice surfaces. Differentiating Open, Transitional, and Frozen Melt Ponds is challenging because the boundaries between these classes are often weak, their optical characteristics partially overlap, and local illumination, ice texture, and surface roughness introduce substantial within-class variability. The overall performance of MP-Unet suggests that its task-oriented architectural components provide complementary benefits under these conditions. Residual blocks with channel attention facilitate the retention and recalibration of subtle spectral and textural information, while attention-gated skip connections reduce the transfer of irrelevant background features into the decoder. The auxiliary binary branch further constrains the network to distinguish melt pond foreground from the surrounding sea-ice surface before resolving finer stage differences. The comparatively poor performance of the ASPP-only configuration indicates that enlarging the receptive field is not sufficient by itself; multi-scale contextual information becomes useful only when combined with effective feature representation and selection. Nevertheless, the lower IoU of Open Melt Ponds relative to the Transitional and Frozen classes shows that class-specific uncertainty

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

remains. This may be related to the smaller representation of Open Melt Ponds in some scenes, their spectral similarity to leads or open-water features, and the variability of their boundaries and internal appearance.

The stage-resolved results reveal substantial heterogeneity in melt pond composition among the five observation sites. Transitional Melt Ponds dominated the pooled pond area and most individual sites, whereas S6 contained a notably larger proportion of Frozen Melt Ponds. Because the observations were obtained during a late-summer period characterized by predominantly subfreezing but near-melting air temperatures, high humidity, and variable wind conditions, the prevalence of transitional surfaces is physically plausible. Under such conditions, pond surfaces may undergo partial refreezing, thinning or breakup of surface ice, and short-term changes in liquid-water exposure. However, the UAV surveys represent spatial snapshots rather than continuous observations of the same ponds. The three mapped classes should therefore be interpreted primarily as observable surface states that are consistent with different stages of melt and refreezing, rather than as a directly observed chronological sequence. Differences among sites may reflect not only the timing of atmospheric forcing but also pond depth, salinity, snow and ice properties, drainage history, floe deformation, and local energy balance. In particular, the contrast between S5 and S6, which were observed on the same date but exhibited markedly different stage compositions, indicates that observation date or latitude alone cannot explain the spatial variability.

The area–frequency distributions consistently show that small and intermediate ponds are numerous, whereas large ponds are comparatively rare. This general pattern agrees with the scale-dependent organization commonly reported for Arctic melt ponds. However, the moderate coefficients of determination and the variation in fitted exponents among sites indicate that the observed distributions should not be interpreted as evidence for a universal power law across the complete area range. Pond coalescence, drainage, fragmentation, segmentation resolution, and the finite spatial extent of the UAV scenes can all alter the shape of the distribution, particularly at the smallest and largest scales. The sparse large-area tail nevertheless remains important because a small number of large ponds can contribute disproportionately to total pond coverage and may exert a correspondingly strong effect on surface albedo. The stage-resolved distributions also show that the areal importance of a class cannot be inferred from pond number alone: a class represented by relatively few objects may still occupy a substantial fraction of the total pond area if its objects are large or connected.

Circularity provides additional evidence that pond surface state is associated with object-level morphology, although the relationship is neither uniform nor necessarily monotonic. Frozen Melt Ponds were generally concentrated toward higher circularity values in the pooled sample, suggesting comparatively compact and regular pond outlines, whereas Open Melt Ponds showed a wider distribution extending toward lower circularity values and therefore greater geometric irregularity. Transitional Melt Ponds commonly occupied an intermediate but broad range. These differences may arise because open water can occupy interconnected depressions and irregular drainage pathways, whereas surface freezing may smooth internal optical contrasts or divide larger pond systems into apparently more compact units. At the same time, the departures observed at individual sites show that morphology cannot be attributed to pond stage alone. Local pond connectivity, ice roughness, drainage history, image resolution, and the way adjacent stage patches are separated during object extraction may all affect the measured perimeter and circularity. Consequently, circularity is most informative when interpreted together with stage composition, area distributions, and local environmental context rather than as an independent indicator of a fixed evolutionary sequence.

The DEM-assisted analysis further indicates that the relationship between melt pond stage and microtopography is expressed more clearly through local relative elevation than through simple ridge proximity. Transitional Melt Ponds occupied significantly lower local positions than Frozen Melt Ponds, whereas the three classes did not differ significantly in their distances to the nearest ridge-like feature. This distinction is important because it suggests that stage differentiation is not simply caused by one class occurring systematically closer to ridges. Instead, ridge-related deformation and surrounding surface relief may generate heterogeneous depressions and drainage configurations within similar ridge-distance ranges. The weak but significant positive relationship between ridge distance and relative elevation for Transitional Melt Ponds is consistent with a tendency for some transitional surfaces near ridge-like features to occupy locally depressed settings. Such depressions may retain meltwater while also experiencing partial surface refreezing under late-season atmospheric conditions. Frozen Melt Ponds, in contrast, occupied relatively higher local positions on average, which may be associated with shallower water, reduced heat storage, or more effective cooling, although these mechanisms cannot be directly resolved from the available data.

The absence of strong relationships for Open and Frozen Melt Ponds and the non-significant overall distance-by-stage interaction require cautious interpretation. The topographic results provide evidence of

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

stage-dependent spatial association, but they do not establish that ridges determine pond stage or drive a unidirectional transition. Ridge-like features were manually interpreted, the DEM represents surface geometry at the time of the survey, and the analysis was conducted at one representative site rather than across all five sites. Furthermore, local relative elevation may itself be affected by photogrammetric uncertainty over low-texture water and ice surfaces. Repeated UAV surveys of the same floes, independently validated elevation data, measurements of pond depth and surface-ice thickness, and pond-scale energy-balance observations would be required to distinguish whether the observed topographic patterns primarily reflect water accumulation, drainage, freezing, or pre-existing ice deformation. Despite these limitations, combining stage-aware segmentation with object-level morphology and relative topographic position provides a more process-oriented observational framework for characterizing melt pond conditions than binary mapping alone. It enables observable surface-state heterogeneity to be linked quantitatively with pond geometry and local terrain, while avoiding unsupported assumptions that spatially separated sites necessarily represent a single temporal sequence.

6 Conclusion

This study developed MP-Unet, a stage-aware semantic segmentation framework for identifying Open, Transitional, and Frozen Melt Ponds from high-resolution UAV imagery. By combining residual channel-attention blocks, multi-scale contextual representation, attention-gated skip connections, and auxiliary foreground supervision, the model achieved an F1-score of 0.9440 and an mIoU of 0.7466 across complex sea-ice surface conditions.

The stage-resolved classification enabled melt pond characteristics to be examined beyond conventional binary pond mapping. Transitional Melt Ponds dominated the pooled pond area, while stage composition varied markedly among observation sites. Pond area distributions showed a general scale-dependent decline with a sparse large-area tail, and the three stages exhibited distinguishable but site-dependent circularity patterns. DEM-assisted analysis further showed that Transitional Melt Ponds occupied lower local topographic positions than Frozen Melt Ponds, although their distances to ridge-like features did not differ significantly. Thus, the most evident ridge-related signal was associated with local topographic position rather than direct control of pond occurrence or planform morphology.

Overall, the study establishes an integrated workflow linking deep-learning-based surface-state

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

classification with object-level morphological and microtopographic analyses. This framework provides a practical approach for characterizing heterogeneous late-summer melt pond conditions at centimetre-scale resolution and may support satellite product validation and the development of more stage-sensitive representations of melt ponds in sea-ice models. Future applications should include repeated observations of the same ice floes, broader spatial validation, and independent measurements of surface elevation and pond thermodynamic properties.

Appendix A

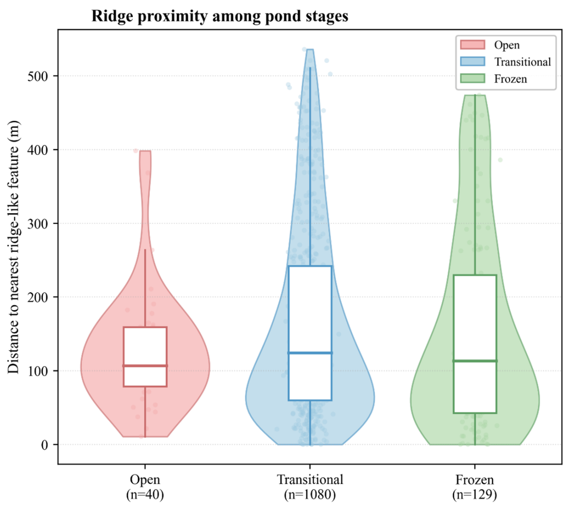


Fig A1. Distribution of ridge proximity among different melt pond stages.

Violin plots combined with boxplots illustrate the distributions of distances from individual melt ponds to the nearest ridge-like feature. Although Transitional and Frozen Melt Ponds exhibit wider distributions than Open Melt Ponds, no statistically significant differences were detected among the three stages (Kruskal–Wallis test, $p > 0.05$).

Code and data availability

The UAV imagery, orthomosaics, DEM data, and derived datasets supporting the findings of this study are available from the corresponding authors upon reasonable request.

The source code used for model training, inference, ablation experiments, and subsequent quantitative analyses is available from the corresponding authors upon reasonable request.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

Author contributions

Wenyu Shen conceived the study, designed the methodology, developed the software, conducted the investigation, performed the formal analysis, prepared the visualisations, and wrote the original draft of the manuscript. Xiaoping Pang and Meng Qu contributed to the conceptualization of the study, supervised the research, and reviewed and edited the manuscript. Guokun Lv contributed to the investigation, data curation, and resources. Xunxiang Wu contributed to the methodology, software development, and validation. Farui Jiang contributed to the visualisations and reviewed and edited the manuscript. Jianbo Shao contributed to data curation and formal analysis. Na Li provided resources and technical support. All authors discussed the results and approved the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank the members of the 14th Chinese National Arctic Research Expedition for logistical support during fieldwork. We also thank the shipboard and UAV operation teams for their assistance with data acquisition.

Financial support.

This work was supported by the National Key Research and Development Program of China, Grant No. 2021YFC2803300 and the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China, Grant No. JYB2025XDXM111.

Reference

Aparício, S., Spreen, G., & Istomina, L. (2023). *Review article: Earth observations of melt ponds on sea ice* [Preprint]. *The Cryosphere Discussions*. https://doi.org/10.5194/tc-2023-75

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). *Encoder-decoder with atrous separable convolution for semantic image segmentation*. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 801–818). Springer. https://doi.org/10.1007/978-3-030-01237-3_69

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

CIE. 2007. Colorimetry — Part 4: CIE 1976 L*a*b* colour space. CIE S 014-4/E:2007 / ISO 11664-4:2008(E). Vienna: International Commission on Illumination.

Ding, Y., Cheng, X., Liu, J., Hui, F., Wang, Z., & Chen, S. (2020). Retrieval of melt pond fraction over Arctic sea ice during 2000–2019 using an ensemble-based deep neural network. *Remote Sensing, 12*(17), 2746. https://doi.org/10.3390/rs12172746

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). IEEE. https://doi.org/10.1109/CVPR.2016.90

Heuzé, C., Jahn, A. The first ice-free day in the Arctic Ocean could occur before 2030. Nat Commun 15, 10101 (2024). https://doi.org/10.1038/s41467-024-54508-3

Hohenegger, C., Alali, B., Steffen, K. R., Perovich, D. K., & Golden, K. M. (2012). Transition in the fractal geometry of Arctic melt ponds. *The Cryosphere, 6*(6), 1157–1162. https://doi.org/10.5194/tc-6-1157-2012

Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. (2020). Squeeze-and-excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 42*(8), 2011–2023. https://doi.org/10.1109/TPAMI.2019.2913372

Huang, W., Lu, P., Lei, R., Xie, H., & Li, Z. (2016). Melt pond distribution and geometry in high Arctic sea ice derived from aerial investigations. *Annals of Glaciology, 57*(73), 105–118. https://doi.org/10.1017/aog.2016.30

Krumpen, T., von Albedyll, L., Goessling, H. F., Hendricks, S., Juhls, B., Spreen, G., Willmes, S., Belter, H. J., Dethloff, K., Haas, C., Kaleschke, L., Katlein, C., Tian-Kunze, X., Ricker, R., Rostosky, P., Rückert, J., Singha, S., & Sokolova, J. (2021). MOSAiC drift expedition from October 2019 to July 2020: Sea ice conditions from space and comparison with previous years. *The Cryosphere, 15*, 3897–3920. https://doi.org/10.5194/tc-15-3897-2021

Lee, S., Stroeve, J., Tsamados, M., & Khan, A. L. (2020). Machine learning approaches to retrieve pan-Arctic melt ponds from visible satellite imagery. *Remote Sensing of Environment, 247*, 111919. https://doi.org/10.1016/j.rse.2020.111919

Lu, P., Leppäranta, M., Cheng, B., Li, Z., Istomina, L., & Heygster, G. (2018). The color of melt ponds on Arctic sea ice. *The Cryosphere, 12*(4), 1331–1345. https://doi.org/10.5194/tc-12-1331-2018

Lund, B., Graber, H. C., Persson, P. O. G., Smith, M., Doble, M., Thomson, J., & Wadhams, P. (2018).

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

Arctic sea ice drift measured by shipboard marine radar. *Journal of Geophysical Research: Oceans, 123*(6), 4298–4321. https://doi.org/10.1029/2018JC013769

Miao, X., Xie, H., Ackley, S. F., Perovich, D. K., & Ke, C. (2015). Object-based detection of Arctic sea ice and melt ponds using high spatial resolution aerial photographs. *Cold Regions Science and Technology, 119*, 211–222. https://doi.org/10.1016/j.coldregions.2015.06.014

Moon, T. A., Druckenmiller, M. L., & Thoman, R. L. (Eds.). (2024). *NOAA Arctic Report Card 2024: Executive summary* (NOAA Technical Report OAR ARC 24-01). NOAA. https://doi.org/10.25923/b7c7-6431

Niehaus, H., Spreen, G., Istomina, L., & Nicolaus, M. (2025). Regional and seasonal evolution of melt ponds on Arctic sea ice. *The Cryosphere, 19*, 3915–3938. https://doi.org/10.5194/tc-19-3915-2025

Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M. C. H., Heinrich, M., Misawa, K., Mori, K., McDonagh, S. G., Hammerla, N. Y., Kainz, B., Glocker, B., & Rueckert, D. (2018). *Attention U-Net: Learning where to look for the pancreas*. *arXiv*. https://arxiv.org/abs/1804.03999

Perovich, D. K., Light, B., Eicken, H., Jones, K. F., Runciman, K., & Nghiem, S. V. (2008). Sunlight, water, and ice: Extreme Arctic sea ice melt during the summer of 2007. *Geophysical Research Letters, 35*(11), L11501. https://doi.org/10.1029/2008GL034007

Polashenski, C., Perovich, D., & Courville, Z. (2012). The mechanisms of sea ice melt pond formation and evolution. *Journal of Geophysical Research: Oceans, 117*, C01001. https://doi.org/10.1029/2011JC007231

Popović, P., Cael, B. B., Silber, M., & Abbot, D. S. (2018). Simple rules govern the patterns of Arctic sea ice melt ponds. *Physical Review Letters, 120*(14), 148701. https://doi.org/10.1103/PhysRevLett.120.148701

Reil, M., Spreen, G., Huntemann, M., Buth, L., & Wilson, D. G. (2024). *Machine learning for the detection of Arctic melt ponds from infrared imagery*. In *Tackling Climate Change with Machine Learning*, ICLR 2024 workshop. https://doi.org/10.48550/arXiv.2404.03379

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional networks for biomedical image segmentation*. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Lecture Notes in Computer Science, Vol. 9351, pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28

Smith, A. R. 1978. Color gamut transform pairs. ACM SIGGRAPH Computer Graphics, 12(3), 12–19.

https://doi.org/10.5194/egusphere-2026-4430
Preprint. Discussion started: 10 August 2026

https://doi.org/10.1145/965139.807361

Stroeve, J., & Notz, D. (2018). Changing state of Arctic sea ice across all seasons. Environmental Research Letters, 13(10), 103001. https://doi.org/10.1088/1748-9326/aade56

Sudakow, I., Asari, V. K., Liu, R., & Demchev, D. (2022). MeltPondNet: A Swin Transformer U-Net for detection of melt ponds on Arctic sea ice. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15*, 8776–8784. https://doi.org/10.1109/JSTARS.2022.3213192

Wang, M., Su, J., Li, T., Wang, X., Ji, Q., Cao, Y., Lin, L., & Liu, Y. (2018). Determination of Arctic melt pond fraction and sea ice roughness from unmanned aerial vehicle (UAV) imagery. *Advances in Polar Science, 29*(3), 181–189. https://doi.org/10.13679/j.advps.2018.3.00181

Webster, M. A., Holland, M., Wright, N. C., Hendricks, S., Hutter, N., Itkin, P., Light, B., Linhardt, F., Perovich, D. K., Raphael, I. A., Smith, M. M., von Albedyll, L., & Zhang, J. (2022). Spatiotemporal evolution of melt ponds on Arctic sea ice: MOSAiC observations and model results. Elementa: Science of the Anthropocene, 10(1), 000072. https://doi.org/10.1525/elementa.2022.00072

Zhou, J., Lu, P., Li, X., Yu, M., Zhang, H., Huo, P., Wang, Q., Shao, H., Li, Z., & Li, Z. (2025). A lightweight machine learning-based melt pond extraction scheme for oblique shipborne imagery. *International Journal of Digital Earth*. https://doi.org/10.1080/17538947.2025.2543573