Dust-Mamba: An Efficient Dust Storm Detection Network with Multiple Data Sources

Cong Bai[1], Zhonghao Lin[1], Jinglin Zhang[2] *, Shengyong Chen[3]

[1] College of Computer Science, Zhejiang University of Technology
[2] School of Control Science and Engineering, Shangdong University
[3] School of Computer Science and Engineering, Tianjin University of Technology
congbai@zjut.edu.cn,201906120421@zjut.edu.cn,jinglin.Zhang@sdu.edu.cn,sy@ieee.org

Abstract

Accurate detection of dust storms is challenging due to complex meteorological interactions. With the development of deep learning, deep neural networks have been increasingly applied to dust storm detection, offering better learning and generalization capabilities compared to traditional physical modeling. However, existing methods face some limitations, leading to performance bottlenecks in dust storm detection. From the task perspective, existing research focuses on occurrence detection while neglecting intensity detection. From the data perspective, existing research fails to explore the utilization of multi-source data. From the model perspective, most models are built on convolutional neural networks, which have an inherent limitation in capturing long-range dependencies. To address the challenges mentioned, this study proposes Dust-Mamba. To the best of our knowledge, this study is the first attempt to accomplish both the occurrence and intensity detection of dust storms with advanced deep learning technology. In Dust-Mamba, multi-source data is introduced to provide a comprehensive perspective, Mamba and attention are applied to boost feature selection while maintaining long-range modeling capability. Additionally, this study proposes Structure Sharing Transfer Learning Strategies for intensity detection, which further enhances the performance of Dust-Mamba with minimal time cost. As shown by experiments, Dust-Mamba achieves Dice scores of 0.963 for occurrence detection and 0.560 for intensity detection, surpassing several baseline models. In conclusion, this study offers valuable baselines for dust storm detection, with significant reference value and promising application potential.

Introduction

Dust storms are a disastrous weather phenomenon that frequently occur in northern and northwestern China (Zhou and Zhang 2003; Wang et al. 2004), West Asia (Hamidi, Kavianpour, and Shao 2013; Khaniabadi et al. 2017) and North Africa (El-ossta, Qahwaji, and Ipson 2013). Severe dust storms have a serious impact on human health and the environment (Goudie 2009; Middleton and Kang 2017), emphasizing the urgent need to improve the precision of dust storm detection.

As shown in Figure 1, compared to existing research, our work aims to use multi-source data to detect both the occurrence and intensity of dust storms, where occurrence detection identifies the presence and intensity detection evaluates intensity levels. Traditional methods employ physical models to detect dust storms, and satellite remote sensing data has been widely utilized (Li et al. 2021). A typical method is to select specific satellite bands and model them for detection. The Infrared Difference Dust Index (IDDI) (Legrand, Plana-Fattori, and N'doumé 2001) is an early practice for dust remote sensing in the Sahel and Sahara regions. The Normalized Difference Dust Index (NDDI) (Qu et al. 2006) utilizes the 0.469 and 2.13 μm bands of Moderate Resolution Imaging Spectroradiometer (MODIS) to detect dust storms, such as in northern China and southern Mongolia. The Brightness Temperature Adjusted Dust Index (BADI) (Yue et al. 2017) uses the brightness temperatures of three thermal infrared MODIS bands ($band_{20}$, $band_{31}$, $band_{32}$) to provide high accuracy detection of dust storms. However, due to the complexity of weather systems, physical models struggle to comprehensively consider all influencing factors, limiting their detection and generalization capabilities (Rivas-Perea, Rosiles, and Chacon 2010; Li et al. 2021).

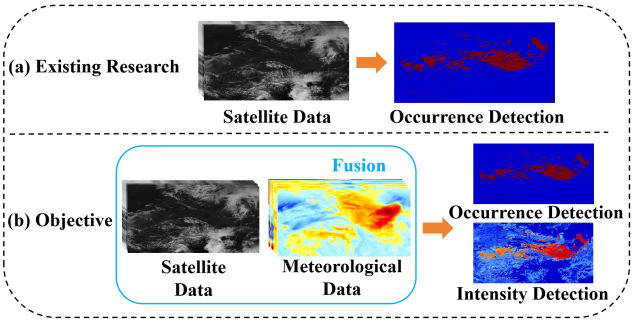

Figure 1: (a) Existing research predominantly utilizes satellite data to detect the occurrence of dust storms. (b) Our research aims to utilize multi-source data as input to detect both the occurrence and intensity of dust storms.

With the development of deep learning, deep neural networks have become powerful tools to analyze complex data (Ma, Pan, and Bai 2024; Zhao et al. 2023). Some studies have explored the application of convolutional neu-

*Corresponding Author

ral networks in dust storm detection tasks. An ensemble model (Bandara 2022) combines U-net (Ronneberger, Fischer, and Brox 2015), Swin-unet (Cao et al. 2022), and Deeplabv3+ (Chen et al. 2018) to detect dust storm regions in MODIS images. The dust storm mask (Jiang et al. 2022) detects dust storms in the Tarim Basin based on Chinese geostationary weather satellites Fengyun-4 A (FY-4 A) (Yang et al. 2017) images. Authors of the Large-Scale Dust Storm Database (LSDSSIMR) (Bai et al. 2023) evaluate deep learning models such as Unet++ (Zhou et al. 2018) and Attention u-net (Oktay et al. 2018) for detecting dust storms. These studies demonstrate that semantic segmentation algorithms are suited for dust storm detection tasks. Additionally, some researchers design Multi-Stream Attention-Aware Convolutional Neural Network (Wang et al. 2023) to detect dust storms with surveillance cameras, and their research provides a valuable supplement to dust storm detection. However, there is potential in model design: pure convolutional architectures focus on local regions and face inherent challenges in capturing long-range dependencies, which are essential for considering distant conditions in dust storm detection.

Based on our observations, current research still faces the following challenges:

a) **Single detection task:** Current studies primarily focus on detecting the occurrence of dust storms, while studies on intensity detection are few. On one hand, occurrence detection is more basic and easier to be implemented, while intensity detection is more challenging. On the other hand, the datasets used in previous research have limitations in this aspect as well.

b) **Lack of multi-source data:** Most studies are based on single-source data, such as satellite data, without integrating multi-source data. In physical modeling, it is challenging to develop equations that effectively integrate multi-source data, while in deep learning, the key problem lies in combining multi-source data with different spatial and temporal resolutions as inputs.

In recent years, Mamba (Gu and Dao 2023) has led to significant breakthroughs in long-range modeling, and VMamba (Liu et al. 2024) extends Mamba's application to computer vision tasks, with the VSS BLOCK as its core module. The Vision Mamba U-net (Ruan and Xiang 2024) that integrates the VSS BLOCK into the U-net architecture achieves excellent results in medical image segmentation tasks. Therefore, Mamba offers a new approach to the design of dust storm detection models.

To address the mentioned challenges, this paper proposes Dust-Mamba which integrates satellite data and meteorological data for dust storm detection tasks. With the application of the proposed CBAM-enhanced VSS BLOCK and Structure Sharing Transfer Learning Strategies, Dust-Mamba improves the accuracy in occurrence and intensity detection. This study demonstrates the effectiveness of multi-source data fusion, long-range modeling, and transfer learning strategies in dust storm detection tasks as well.

The main contributions of our work are summarized below:

1) **Multi-source Data Fusion:** To combine data from different sources, we propose Multi-Resolution Data Fusion (MRDF). Features from meteorological data are processed independently during the feature extraction stages. Then meteorological features are integrated into processed fused features at the same scale during the fusion processes. This reduces the impact of different temporal and spatial resolutions and fully exploits the information from different data sources.
2) **Long-range Modeling with Attention:** To overcome the limitations of convolutional neural networks and capture long-range dependencies, we propose CBAM-enhanced VSS BLOCK. The features are processed in two parallel branches, where global features are extracted through the VSS BLOCK and attention-weighted features are captured by the Convolutional Block Attention Module (CBAM). By merging the outputs from two branches, it enables Dust-Mamba to extract global features and selectively focus on them.
3) **Transfer Learning Strategy:** To improve intensity detection, we propose Structure Sharing Transfer Learning Strategies, which markedly outperform models trained from scratch with minimal time cost. By fine-tuning a model trained for occurrence detection, we effectively improve the precision and reduce training time. By training occurrence detection and intensity detection models together, we achieve substantial performance gains with a reasonable time increase.

Method

Problem Definition

Both occurrence detection and intensity detection require pixel-wise classification, which can be seen as semantic segmentation tasks (Lateef and Ruichek 2019; Hao, Zhou, and Guo 2020; Yuan, Shi, and Gu 2021). We define the model as $Y = \mathit{model}(X)$. X is input data that consists of multi-channel information with size $B \times C \times W \times H$, and Y is output data with size $B \times S \times W \times H$. B represents the batch size, C denotes the number of channels in the input data, W and H refer to the width and height of the data matrix. S represents the output channels, where $S = 2$ for dust storm occurrence detection, indicating the presence or absence of a dust storm, and $S = 7$ for intensity detection, denoting that the model outputs seven levels of intensity.

Overall Architecture

The overall architecture of our Dust-Mamba is shown in Figure 2(a). Unet++ is selected as the main structure due to its nested and dense skip connections, which effectively improve detection accuracy. In the encoding stage, the input data is divided into satellite data and meteorological data. The satellite data is fused with meteorological data, then they are processed through four MRDF modules and a convolutional module to obtain encoded features. In the decoding stage, the encoded features are gradually restored to the size $W \times H$ through four decoder layers. Each decoder involves nested and dense skip connections, along with an upsampling module and a convolutional module. In the output

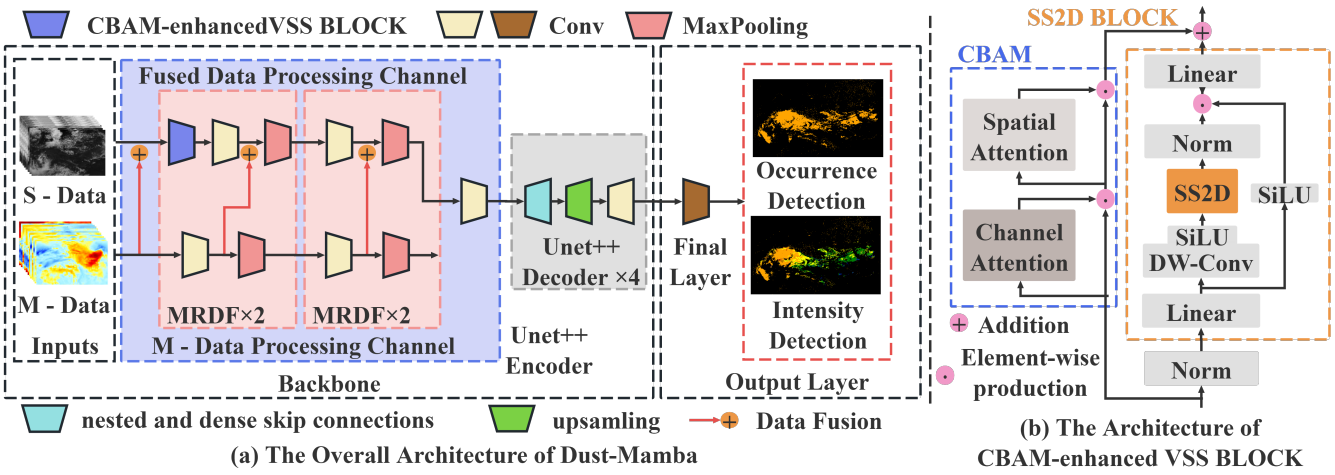

Figure 2: The Overall Architecture of Dust-Mamba and Our CBAM-enhanced VSS BLOCK

stage, the convolutional layer produces detection results depending on the specific task.

Multi-Resolution Data Fusion

There are inherent differences between multi-source data. For instance, within an hour, meteorological data remains constant, whereas satellite data updates four times. We find that concatenating all data only at the input stage is not a wise choice. To achieve better feature extraction, we adopt Multi-Resolution Data Fusion (MRDF), processing satellite and meteorological data with relative independence. Compared to meteorological data, satellite data reflect dust storm information more accurately, so we consider satellite data as the primary data, with meteorological data serving as auxiliary information. In the MRDF, the meteorological data is processed independently through convolutional modules to extract multi-scale features. Then, a series of meteorological features is integrated into the processed fused features of the same scale and further processed through subsequent modules such as our CBAM-enhanced VSS BLOCK.

CBAM-enhanced VSS BLOCK

Recently, State Space Models (SSMs) garner widespread research interest. Building on classical SSM (Kalman 1960), modern SSM-based models like Mamba (Gu and Dao 2023) exhibit excellent performance in long-range modeling. S6 models (Gu and Dao 2023), which are derived from the Structured State Space Sequence models (S4) (Gu, Goel, and Ré 2021), introduce dynamic weights and adjust the SSM parameters based on the inputs, forming the foundation of the Mamba. VMamba (Liu et al. 2024), a visual state space model, extends the applicability of the Mamba from one-dimensional data to two-dimensional data, where the VSS BLOCK is the core, and the SS2D BLOCK is the core of the VSS BLOCK. The SS2D BLOCK primarily achieves expanding, S6-based modeling and merging, allowing each pixel to capture information from other pixels, ensuring long-range modeling capability. The visualization

Algorithm 1: S6 BLOCK in SS2D

Input: x (batch size, length, dimension)
Output: y (batch size, length, dimension)

$\Delta, \mathbf{B}, \mathbf{C} = \text{Linear}(x), \text{Linear}(x), \text{Linear}(x)$
$\overline{\mathbf{A}}, \overline{\mathbf{B}} = \exp(\Delta\mathbf{A}), (\Delta\mathbf{A})^{-1}\left(\exp(\Delta\mathbf{A}) - \mathbf{I}\right) \cdot \Delta\mathbf{B}$
$h_t = \overline{\mathbf{A}} h_{t-1} + \overline{\mathbf{B}} x_t$
$y_t = \mathbf{C} h_t + \mathbf{D} x_t$
$y = [y_1, y_2, \cdots, y_t, \cdots, y_L]$
return y

of the SS2D is shown in Figure 3 and the pseudo-code of the S6 BLOCK is shown in Algorithm 1. VM-UNet (Ruan and Xiang 2024) is an application of Mamba in semantic segmentation tasks. It incorporates VSS BLOCK into the U-net structure, demonstrating the effectiveness of long-range modeling in medical image segmentation tasks and offering valuable insights for other domains.

Dust-Mamba integrates the VSS BLOCK to leverage its long-range modeling advantages as well. We recognize that: (1) The data used in our study differs from typical 3-channel images, as it consists of multiple channels of satellite and meteorological data. (2) Different satellite and meteorological channels have distinct impacts on our tasks due to their various characteristics. Thus, CBAM (Woo et al. 2018), a hybrid attention mechanism that combines channel and spatial attention modules, is incorporated into VSS BLOCK. The design is shown in Figure 2(b), the CBAM is added into the residual connection pathway. By employing CBAM and SS2D in parallel, they can independently extract features from the original inputs, preventing feature distortion that may arise from sequential processing. If we view SS2D as a specialized attention, the CBAM-enhanced VSS BLOCK enables Dust-Mamba to capture both long-range dependencies and local details, where VSS BLOCK focuses on long-range dependencies, channel attention prioritizes the relevant features, and spatial attention emphasizes critical local regions within each feature map.

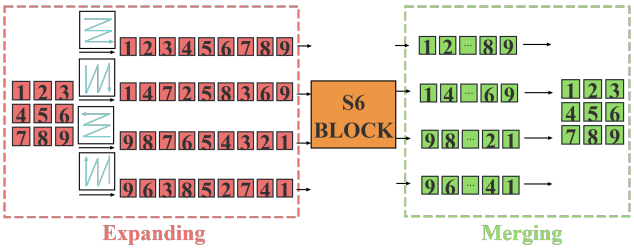

Figure 3: The Visualization of SS2D BLOCK. *Four expansion directions : top-left to bottom-right and bottom-right to top-left, along with horizontal and vertical.*

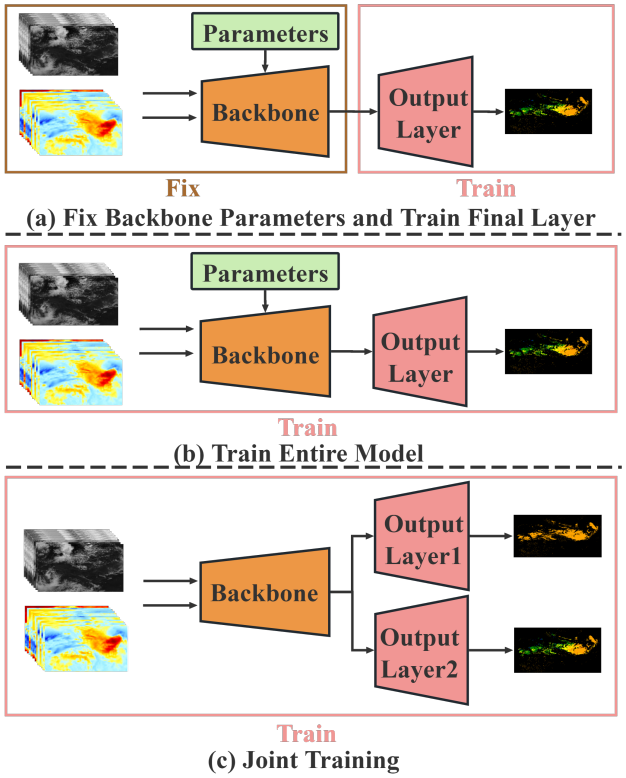

Figure 4: The Structure Sharing Transfer Learning Strategies. *Strategies (a) and (b) introduce the parameters from the occurrence detection model. (c) adopts two outputs.*

Structure Sharing Transfer Learning Strategy

Dust-Mamba has the same input data and main structure for both detection tasks, but differs in the final layer. Inspired by transfer learning (Ribani and Marengoni 2019; Zhuang et al. 2020; Lin et al. 2024) in accelerating training process, we propose Structure Sharing Transfer Learning Strategies for intensity detection in two main directions. (1) Introducing the backbone parameters of the occurrence detection model and then fine-tuning the intensity detection model, which can be further divided into fine-tuning the final output layer or the entire model. (2) Jointly training the occurrence detection and intensity detection models, with each task having a separate output. These strategies are shown in Figure 4.

Experiments

Dataset

LSDSSIMR (Bai et al. 2023) covers dust storms from March to May each year between 2020 and 2022, in the region between 38.9°E to 153.8°E longitude and 20.0°N to 58.2°N latitude. It provides a temporal resolution of 15 minutes and a spatial resolution of 4 km. We divide the dataset into a training set (2020 - 2021), a validation set (March 2022), and a test set (April and May 2022). After selection, experiments include 3923 samples for training, 440 for validation, and 1,045 for testing, with a ratio of approximately 7:1:2. The satellite and meteorological data serve as input data, while the DST data serves as label data, as shown in Table 1.

Baselines and Evaluation Metric

Following semantic segmentation models are compared. U-net (Ronneberger, Fischer, and Brox 2015), a classic segmentation network, is originally designed for medical image segmentation. Unet++ (Zhou et al. 2018) incorporates dense convolution blocks in the skip connections. Attention u-net (Oktay et al. 2018) focuses on target areas by embedding attention gates in the skip connections. Transunet (Chen et al. 2021) combines CNNs and Transformer in the encoder network. Swin-unet (Cao et al. 2022) and VM-unet (Liu et al. 2024) respectively introduce Swin Transformer BLOCK and VSS BLOCK into U-net architecture, replacing traditional convolutional layers. The evaluation metrics include Dice, Iou, Recall, Precision, Kappa (Bai et al. 2023) and higher values indicate superior performance.

Implementation Details

All experiments run on the NVIDIA A6000 GPUs are repeated three times. Each model is trained for 200 epochs, and the best-performing one is selected based on the validation set. To fit the model, we crop the original data size from 640×1280 to 448×896. The AdamW optimizer is employed, starting with a dynamic decay of 5e-4. The loss function combines Dice loss and Cross-Entropy loss, defined as $0.5 \times$ Dice loss $+ 0.5 \times$ Cross-Entropy loss. The drop rate is set to 0.1.

Results and Analyses

Ablation Study

To validate the effectiveness of each component of Dust-Mamba, we conduct experiments by adding modules step by step from the original Unet++. The ablation study results for the different modules are shown in Table 2, with the best results highlighted in bold (the same applies to the following tables). The top half covers occurrence detection ablation, and the bottom half covers intensity detection ablation. The first rows in both parts show the performance of the original Unet++. The second rows indicate MRDF improves the Dice scores by 1.5% in occurrence detection and 1.5% in intensity detection. Therefore, it can be concluded that multi-source data is integrated effectively by MRDF, ensuring the effective extraction of features from different temporal and spatial resolutions. The third rows represent VSS BLOCK

Data Type	Dataset Fields	Unit	Data Source	Tem-Resolution	Spa-Resolution
Input Satellite Data	R0.47	K	FY-4A	5 min	4km×4km
	R0.65	K	FY-4A	5 min	4km×4km
	BT3.72	K	FY-4A	5 min	4km×4km
	BT8.50	K	FY-4A	5 min	4km×4km
	BT10.8	K	FY-4A	5 min	4km×4km
	BT12.0	K	FY-4A	5 min	4km×4km
Input Meteorological Data	surface pressure	hpa	ERA5	1 hour	0.25°× 0.25°
	mean sea level pressure	hpa	ERA5	1 hour	0.25°× 0.25°
	skin temperature	K	ERA5	1 hour	0.25°× 0.25°
	2m air temperature	K	ERA5	1 hour	0.25°× 0.25°
	2m dewpoint temperature	K	ERA5	1 hour	0.25°× 0.25°
	10m air temperature	K	MERRA-2	1 hour	0.625°× 0.5°
	2m eastward wind	m/s	MERRA-2	1 hour	0.625°× 0.5°
	2m northward wind	m/s	MERRA-2	1 hour	0.625°× 0.5°
	10m eastward wind	m/s	ERA5	1 hour	0.25°× 0.25°
	10m northward wind	m/s	ERA5	1 hour	0.25°× 0.25°
	50m eastward wind	m/s	MERRA-2	1 hour	0.625°× 0.5°
	50m northward wind	m/s	MERRA-2	1 hour	0.625°× 0.5°
Label Data	DST	–	FY-4A	5 min	4km×4km

Table 1: The Data of the LSDSSIMR Used in Our Study. *(a) °indicates latitude and longitude. (b) For eastward and northward wind components in the table, positive values indicate the wind is blowing eastward and northward, and negative values indicate the wind is blowing westward and southward.*

Model	Dice	Iou	Rec	Pre	Kap
Unet++(O)	0.9132	0.8444	0.9038	0.9247	0.9118
+MRDF	0.9286	0.8699	0.9251	0.9337	0.9275
+VSB	0.9562	0.9192	0.9415	**0.9737**	0.9555
+CBAM	**0.9637**	**0.9324**	**0.9663**	0.9627	**0.9631**
Unet++(I)	0.4581	0.3499	0.4306	0.6242	0.4701
+MRDF	0.4735	0.3633	0.4490	0.6242	0.4851
+VSB	0.5099	0.3950	0.4921	0.6272	0.5208
+CBAM	**0.5158**	**0.4010**	**0.4939**	**0.6316**	**0.5266**

Table 2: The Ablation Experiment for Dust Storm Occurrence and Intensity Detection. *(O) represents dust storm occurrence detection. (I) represents dust storm intensity detection. Each model builds upon the previous one and the final rows in both parts represent our Dust-Mamba.*

improves the Dice scores by 2.8% in occurrence detection and 3.6% in intensity detection. The results show long-range modeling capability is essential to dust storm detection, and Dust-Mamba leverages the advantages of it to enhance the performance. The fourth rows show CBAM improves the Dice scores by 0.75% in occurrence detection and 0.6% in intensity detection. It demonstrates Dust-Mamba's ability to focus on dust storm-related features and spatial regions, providing valuable support for long-range modeling.

Discussion on Transfer Learning Strategy

In Table 3, different transfer training strategies of the Dust-Mamba in intensity detection are compared. The first row means Dust-Mamba trains from scratch and the following

Strategy	Dice	Iou	Rec	Pre	Kap
Strategy0	0.5158	0.4010	0.4939	0.6316	0.5266
Strategy1	0.2743	0.1991	0.3320	0.3458	0.2810
Strategy2	**0.5604**	**0.4448**	**0.5414**	0.6654	**0.5707**
Strategy3	0.5529	0.4360	0.5250	**0.6683**	0.5631

Table 3: The Different Transfer Learning Strategies of Dust-Mamba in Intensity Detection. *Strategy0: train from scratch, Strategy1: train final layer, Strategy2: train entire model, Strategy3: joint training.*

lines mean our proposed strategies. The second row shows training the final output layer gets unsatisfactory results. In the third and fourth rows, fixing backbone parameters and training the entire model, joint training both effectively improve performance, with these strategies getting 4.5% and 3.7% improvement in terms of the Dice metric. Compared to a model training from scratch, these strategies show strong training efficiency: when training the entire model, it reaches the performance of a model trained from scratch in approximately 1/3 of the time, and when adopting joint training, the model only requires approximately 5% more time per epoch. We believe that intensity detection is inherently linked to occurrence detection, as intensity levels can only be measured after confirming the dust storm occurrence. Training the entire model and joint training leverage occurrence detection model parameters as prior knowledge that can be fine-tuned for intensity detection, enabling the model to better learn the relationship between variables and dust storm intensity levels. As a comparison, simply training the final output layer yields worse results, as backbone parameters are trained to

view dust storms from an occurrence perspective and do not directly apply to intensity detection.

Comparison with Baseline methods

We compare various baselines for occurrence and intensity detection. Two approaches are employed for occurrence detection: the first approach uses only satellite data as input data, specifically utilizing six channels of satellite data, as shown in Table 4, the second approach involves a simple concatenation of the satellite data with meteorological data, incorporating twelve channels of meteorological data alongside the six channels of satellite data, as shown in Table 5. We find that (1) Unet++ delivers the best performance due to the effectiveness of nested and dense skip connections. (2) Simply concatenating satellite and meteorological data as inputs does not markedly enhance the performance and may even degrade it, as the model struggles to distinguish between data of different resolutions. Compare of baselines for intensity detection is shown in Table 6, and the experimental results align with occurrence detection, Unet++ performs the best while Transunet performs the worst.

Model(6)	Dice	Iou	Rec	Pre	Kap
U-net	0.8945	0.8145	0.8981	0.8924	0.8928
Unet++	**0.9134**	**0.8447**	0.8971	**0.9321**	**0.9120**
Attunet	0.8963	0.8180	**0.9005**	0.8935	0.8947
Transunet	0.5316	0.3866	0.4725	0.6604	0.5251
Swinunet	0.8190	0.7025	0.7922	0.8554	0.8161
VMunet	0.8257	0.7111	0.8259	0.8426	0.8227

Table 4: The Performance Comparison for Dust Storm Occurrence Detection with Satellite Data. *(6) indicates that 6 satellite data channels are concatenated as input data.*

Model(18)	Dice	Iou	Rec	Pre	Kap
U-net	0.8806	0.7937	0.8701	0.8943	0.8787
Unet++	**0.9132**	**0.8444**	**0.9038**	**0.9247**	**0.9118**
Attunet	0.8905	0.8090	0.8782	0.9058	0.8888
Transunet	0.4603	0.3225	0.3783	0.6674	0.4536
Swinunet	0.6561	0.5055	0.5816	0.7772	0.6512
VMunet	0.7847	0.6562	0.7380	0.8575	0.7813

Table 5: The Performance Comparison for Dust Storm Occurrence Detection with Satellite Data and Meteorological Data. *(18) indicates that 6 satellite data channels and 12 meteorological data channels are concatenated as input data.*

Then, we compare Dust-Mamba with the baseline models, as shown in Table 7. It presents the comparison between Dust-Mamba and the best-performing baseline, Unet++. The top half of the table shows the occurrence detection results, showing that Dust-Mamba achieves a 5% improvement in the Dice and an 8.8% enhancement in the Iou. The bottom half displays the intensity detection results, showing that our model achieves a 10% improvement in the Dice and a 9.5% enhancement in Iou. We believe Dust-Mamba achieves strong performance in both dust storm occurrence detection and intensity detection.

Model	Dice	Iou	Rec	Pre	Kap
U-net	0.4320	0.3275	0.4081	0.6103	0.4442
Unet++	**0.4581**	**0.3499**	**0.4306**	0.6242	**0.4701**
Attunet	0.4421	0.3355	0.4168	**0.6323**	0.4542
Transunet	0.1482	0.0972	0.1347	0.5283	0.1609
Swinunet	0.2369	0.1561	0.2050	0.5300	0.2495
VMunet	0.3195	0.2196	0.2828	0.5355	0.3317

Table 6: The Performance Comparison for Dust Storm Intensity Detection. *All models concatenate 6 satellite data channels and 12 meteorological data as input data.*

Model	Dice	Iou	Rec	Pre	Kap
Unet++(O)	0.9132	0.8444	0.9038	0.9247	0.9118
Ours(O)	**0.9637**	**0.9324**	**0.9663**	**0.9627**	**0.9631**
Unet++(I)	0.4581	0.3499	0.4306	0.6242	0.4701
Ours(I)	**0.5604**	**0.4448**	**0.5414**	**0.6654**	**0.5707**

Table 7: The Comparison of Dust-Mamba and Baseline Models. *Ours(O) and (I) mean Dust-Mamba trained for occurrence and intensity detection respectively. Ours(I) introduces the parameters from Ours(O) and trains entire model.*

model	D1	D2	D3	D4	D5	D6
U-net	0.1108	0.1960	0.6624	0.6468	0.7296	0.2462
Unet++	0.1310	0.2536	0.7004	0.6759	0.7360	0.2519
Attunet	0.1139	0.1977	0.6587	0.6595	0.7342	0.2887
Transunet	0.0826	0.0241	0.3055	0.2904	0.1631	0.0236
Swinunet	0.0790	0.0730	0.4285	0.3837	0.3138	0.1432
VMunet	0.0880	0.1347	0.4962	0.4568	0.5398	0.2016
Ours	**0.1926**	**0.3743**	**0.7763**	**0.7465**	**0.8456**	**0.4271**

Table 8: The Performance Comparison for Different Dust Storm Intensity Levels. *Ours refers to the Dust-Mamba that introduces parameters and trains the entire model. D1-D6 represent dust storm intensity levels from 1 to 6.*

For intensity detection, we further provide a comparison of various intensity levels, as shown in Table 8. Taking the Dice as examples, all models show weaker detection capabilities for intensity levels 1 and 6. Based on our analysis, in terms of data, levels 1 and 6 comprise only 2.10% and 1.98% of all dust storm levels (excluding level 0), resulting in limited training data. In terms of features, level 1 resembles non-dust-storm conditions, while level 6 differs to some extent from other levels, which hinders the model from fully learning these features. Nevertheless, by applying the proposed methods, Dust-Mamba effectively improves the detection capability for these intensity levels across all metrics, even with limited data.

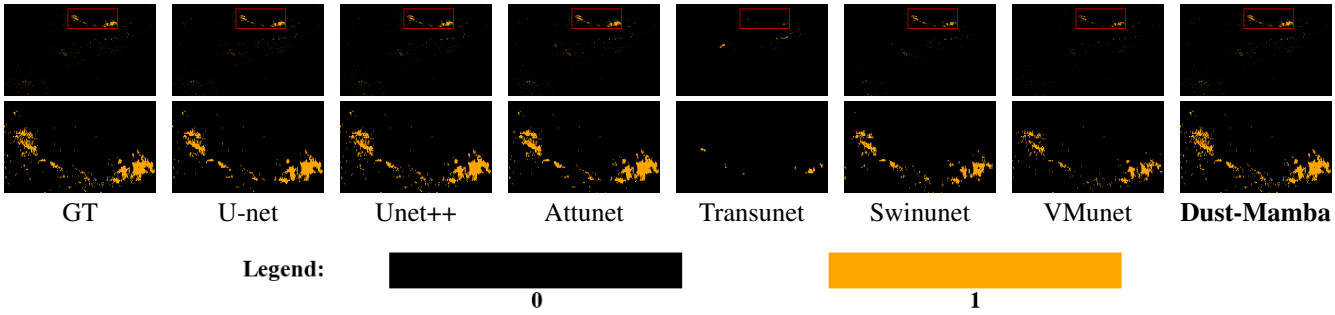

Figure 5: The Visualization for Occurrence Detection. *0 represents no dust storm and 1 represents the presence of a dust storm.*

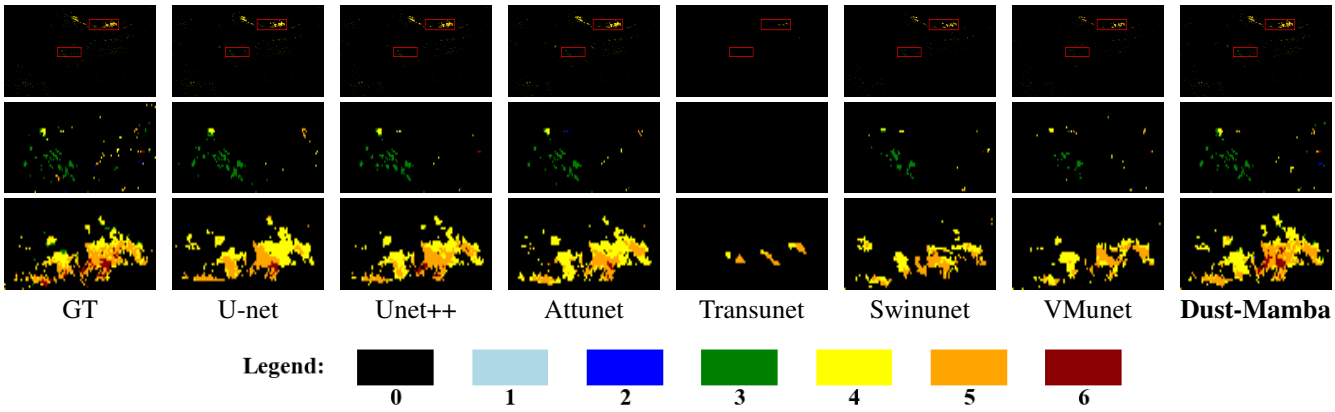

Figure 6: The Visualization for Intensity Detection. *0 represents no dust storm and 1-6 represent dust storm intensity levels.*

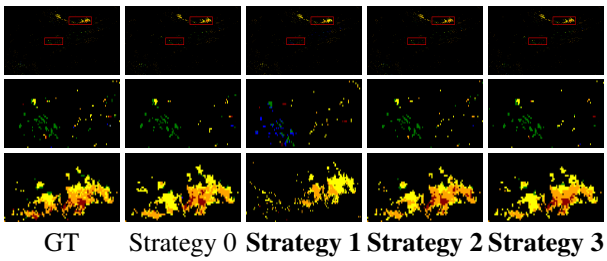

Figure 7: The Visualization for Transfer Learning Strategies of Dust-Mamba. *The legend is same as Figure 6. Strategy 0: train from scratch, Strategy 1: train final layer, Strategy 2: train entire model, Strategy 3: joint training.*

Visualization

We visualize the detection results of all models, as shown in Figure 5, Figure 6 and Figure 7. For every visualization, the first row shows the full results, while the latter rows provide close-ups. For occurrence detection, as shown in Figure 5, Dust-Mamba shows a better fit to the ground truth (GT) and excels in handling finer details compared to well-performing baseline models such as Unet++. For intensity detection, as shown in Figure 6, the second row shows Dust-Mamba detects more dust storms accurately than the baseline models. The third row illustrates that Dust-Mamba closely matches the ground truth in detecting the extent of strong dust storms at level 6, where the baseline models struggle. In Figure 7, the visualizations of different strategies confirm that training only the final output layer results in poorer performance. In contrast, the other two strategies provide more accurate intensity detection, demonstrating their improved detection performance compared to the model trained from scratch.

Conclusion

In this study, we propose Dust-Mamba to address limitations in existing dust storm research, including occurrence detection and intensity detection. Specifically, we propose the MRDF, which integrates satellite and meteorological data to overcome challenges posed by different resolutions. We design the CBAM-enhanced VSS BLOCK, leveraging its strengths in long-range modeling and feature selection. For the intensity detection task, we develop the Structure Sharing Transfer Learning Strategies, enabling the model to learn features more efficiently and accurately. Experimental results show that Dust-Mamba successfully detects the occurrence and intensity of dust storms, surpassing several baseline models. In the future, we plan to enhance detection performance by the following methods: (1) Expanding the dataset for generalization; (2) Designing better model architectures by incorporating physical knowledge; (3) Identifying optimal data input through sensitivity analysis. We believe our research offers valuable insights for future deep learning applications in dust storm detection.

Acknowledgments

This work is partially supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LRG25F020002 and LR21F020002. Code is available at https://github.com/Zjut-MultimediaPlus/Dust-Mamba.

References

Bai, C.; Cai, Z.; Yin, X.; and Zhang, J. 2023. LSDSSIMR: Large-Scale Dust Storm Database Based on Satellite Images and Meteorological Reanalysis Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Bandara, N. 2022. Ensemble Deep Learning for Automated Dust Storm Detection Using Satellite Images. In *2022 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, volume 5, 178–183. IEEE.

Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; and Wang, M. 2022. Swin-unet: Unet-like pure transformer for medical image segmentation. In *European conference on computer vision*, 205–218. Springer.

Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A. L.; and Zhou, Y. 2021. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*.

Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; and Adam, H. 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, 801–818.

El-ossta, E.; Qahwaji, R.; and Ipson, S. S. 2013. Detection of dust storms using MODIS reflective and emissive bands. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(6): 2480–2485.

Goudie, A. S. 2009. Dust storms: Recent developments. *Journal of environmental management*, 90(1): 89–94.

Gu, A.; and Dao, T. 2023. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*.

Gu, A.; Goel, K.; and Ré, C. 2021. Efficiently modeling long sequences with structured state spaces. *arXiv preprint arXiv:2111.00396*.

Hamidi, M.; Kavianpour, M. R.; and Shao, Y. 2013. Synoptic analysis of dust storms in the Middle East. *Asia-Pacific Journal of atmospheric sciences*, 49: 279–286.

Hao, S.; Zhou, Y.; and Guo, Y. 2020. A brief survey on semantic segmentation with deep learning. *Neurocomputing*, 406: 302–321.

Jiang, H.; He, Q.; Zhang, J.; Tang, Y.; Chen, C.; Lv, X.; Zhang, Y.; and Liu, Z. 2022. Dust storm detection of a convolutional neural network and a physical algorithm based on FY-4A satellite data. *Advances in Space Research*, 69(12): 4288–4306.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems.

Khaniabadi, Y. O.; Daryanoosh, S. M.; Amrane, A.; Polosa, R.; Hopke, P. K.; Goudarzi, G.; Mohammadi, M. J.; Sicard, P.; and Armin, H. 2017. Impact of Middle Eastern Dust storms on human health. *Atmospheric pollution research*, 8(4): 606–613.

Lateef, F.; and Ruichek, Y. 2019. Survey on semantic segmentation using deep learning techniques. *Neurocomputing*, 338: 321–348.

Legrand, M.; Plana-Fattori, A.; and N'doumé, C. 2001. Satellite detection of dust using the IR imagery of Meteosat: 1. Infrared difference dust index. *Journal of Geophysical Research: Atmospheres*, 106(D16): 18251–18274.

Li, J.; Wong, M. S.; Lee, K. H.; Nichol, J.; and Chan, P. 2021. Review of dust storm detection algorithms for multispectral satellite sensors. *Atmospheric Research*, 250: 105398.

Lin, F.; Hu, W.; Wang, Y.; Tian, Y.; Lu, G.; Chen, F.; Xu, Y.; and Wang, X. 2024. Universal Object Detection with Large Vision Model. *International Journal of Computer Vision*, 132(4): 1258–1276.

Liu, Y.; Tian, Y.; Zhao, Y.; Yu, H.; Xie, L.; Wang, Y.; Ye, Q.; and Liu, Y. 2024. VMamba: Visual State Space Model. *CoRR*, abs/2401.10166.

Ma, Q.; Pan, J.; and Bai, C. 2024. Direction-Oriented Visual–Semantic Embedding Model for Remote Sensing Image–Text Retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1–14.

Middleton, N.; and Kang, U. 2017. Sand and dust storms: Impact mitigation. *Sustainability*, 9(6): 1053.

Oktay, O.; Schlemper, J.; Folgoc, L. L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N. Y.; Kainz, B.; et al. 2018. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*.

Qu, J. J.; Hao, X.; Kafatos, M.; and Wang, L. 2006. Asian dust storm monitoring combining Terra and Aqua MODIS SRB measurements. *IEEE Geoscience and remote sensing letters*, 3(4): 484–486.

Ribani, R.; and Marengoni, M. 2019. A survey of transfer learning for convolutional neural networks. In *2019 32nd SIBGRAPI conference on graphics, patterns and images tutorials (SIBGRAPI-T)*, 47–57. IEEE.

Rivas-Perea, P.; Rosiles, J.; and Chacon, M. 2010. Traditional and neural probabilistic multispectral image processing for the dust aerosol detection problem. In *2010 IEEE Southwest Symposium on Image Analysis & Interpretation (SSIAI)*, 169–172. IEEE.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, 234–241. Springer.

Ruan, J.; and Xiang, S. 2024. Vm-unet: Vision mamba unet for medical image segmentation. *arXiv preprint arXiv:2402.02491*.

Wang, X.; Dong, Z.; Zhang, J.; and Liu, L. 2004. Modern dust storms in China: an overview. *Journal of Arid Environments*, 58(4): 559–574.

Wang, X.; Yang, Z.; Feng, H.; Zhao, J.; Shi, S.; and Cheng, L. 2023. A Multi-Stream Attention-Aware Convolutional Neural Network: Monitoring of Sand and Dust Storms from Ordinary Urban Surveillance Cameras. *Remote Sensing*, 15(21): 5227.

Woo, S.; Park, J.; Lee, J.-Y.; and Kweon, I. S. 2018. Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, 3–19.

Yang, J.; Zhang, Z.; Wei, C.; Lu, F.; and Guo, Q. 2017. Introducing the new generation of Chinese geostationary weather satellites, Fengyun-4. *Bulletin of the American Meteorological Society*, 98(8): 1637–1658.

Yuan, X.; Shi, J.; and Gu, L. 2021. A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Systems with Applications*, 169: 114417.

Yue, H.; He, C.; Zhao, Y.; Ma, Q.; and Zhang, Q. 2017. The brightness temperature adjusted dust index: An improved approach to detect dust storms using MODIS imagery. *International journal of applied earth observation and geoinformation*, 57: 166–176.

Zhao, D.; Wang, Q.; Zhang, J.; and Bai, C. 2023. Mine diversified contents of multispectral cloud images along with geographical information for multilabel classification. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1–15.

Zhou, Z.; Rahman Siddiquee, M. M.; Tajbakhsh, N.; and Liang, J. 2018. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*, 3–11. Springer.

Zhou, Z.; and Zhang, G. 2003. Typical severe dust storms in northern China during 1954—2002. *Chinese Science Bulletin*, 48: 2366–2370.

Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; and He, Q. 2020. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1): 43–76.