



Adaptive enhancement of spatial information in adverse weather

Mohammad Shabaz¹ · Mukesh Soni^{2,3}

Received: 8 January 2024 / Revised: 15 February 2024 / Accepted: 15 February 2024 / Published online: 21 February 2024
© The Author(s), under exclusive licence to Korea Spatial Information Society 2024

Abstract

In the context of spatial information, particularly in video surveillance and intelligent transportation systems, the visibility of video images is severely impacted by adverse climates including rain, snow, and fog. Accurate and swift recognition of current weather conditions and adaptive clarification of surveillance videos are crucial to maintaining the integrity of spatial information. Addressing the limitations of traditional weather recognition methods and the scarcity of weather image datasets, a multicategory weather image block dataset was constructed. This research introduced a weather recognition algorithm that integrates image block processing with feature fusion. The algorithm uses traditional methods to extract shallow spatial features such as average gradient, contrast, saturation, and dark channel from weather images. It also employs transfer learning to fine-tune a pretrained VGG16 model, extracting deep spatial features from the model's fully connected layers. The approach improves the SoftMax classifier's recognition of fog, rain, snow, and clear weather photos by merging shallow and deep spatial information. This improvement is essential for the quality and reliability of spatial data in bad weather. The algorithm achieves 99.26% accuracy in weather detection; however, the best accuracy archive by state of art is 97.14%, confirming its usefulness as a module for adaptive video picture clarification in spatially informed systems.

Keywords Spatial information · Adverse Weather · Image recognition · Fusion-based Algorithm · Multicategory Weather dataset · Video Surveillance · Intelligent Transportation

1 Introduction

Most computer vision research uses good weather photos or movies. Fog, rain, and snow reduce video visibility, making all-weather monitoring difficult. In outdoor video surveillance, it is crucial to swiftly gather weather conditions and quality video pictures. Conventional weather identification involves sensors or radars [1], but it is restricted and

expensive. Most weather identification research uses cameras to capture footage and computers to identify weather. The three main types of image-based weather recognition algorithms are those that train classifiers using weather image features, those that use deep learning, and those that use fusion features.

Using spatial information from weather photos, a weather identification classifier may be trained. This approach requires spatially-based weather image characteristics. Dictionary learning and multicore learning were used to classify weather images. This approach extracts the characteristics of the picture using spatial data. These features are the sky, shadow, rain line, snowflake, dark channel, contrast, and saturation. By employing a multicore learning algorithm, the optimal weight of all these features is learned. This weight is then used to identify and classify four types of weather images: fog, rain, snow, and clear. According to the source [2], the accuracy rate is reported to be 71%. The traditional feature extraction method, which involves mining, is a spatially oriented method. The extracted features, when analyzed spatially, can effectively capture the

✉ Mohammad Shabaz
shabaz.cse@mietjammu.in

Mukesh Soni
mukesh.research24@gmail.com

¹ Model Institute of Engineering and Technology, Jammu, J&K, India

² Dr. D. Y. Patil Vidyapeeth, Pune, Dr. D. Y. Patil School of Science & Technology, Tathawade, Pune, India

³ Department of CSE, University Centre for Research & Development, Chandigarh University, Mohali 140413, Punjab, India

surface-level information present in the image. However, they do not adequately capture the intricate spatial relationships and underlying semantic meaning embedded within the image. Therefore, the recognition effect of the first type of method is usually poor in terms of spatial difference.

Second, end-to-end recognition based on deep learning. This study created the RFS (Rain Fog Snow) data collection of rain, fog and snow photographs and recommended analysing it with Superpixel masks. Although RFS data collection is private, enhancement and 10 types of convolutional neural networks for training can achieve recognition accuracy greater than 70% [3]. This article labels nine weather photographs. Integrating deep residual and dense convolution networks yields a weather recognition accuracy of 80% [4]. Fang et al. suggested an updated SqueezeNet-based weather classification approach that reduces the parameters to AlexNet equivalent in 1/50 of the original condition, but performs worse than ResNet and VGG16 [5]. This work suggested a block image and voting strategy-based classification approach for outside transmission line weather picture data, including foggy, rainy, snowy and sunny days. Refine ResNet50 to extract weather. The method identifies weather effectively, but lacks shallow weather image processing. Expression of layer data [6]. The characteristics of the deep neural network from end to end express the abstractness of the image and the intrinsic information [7]. The second method improves the accuracy of recognition when trained on large weather image data.

The third type of method combines the first two types of methods by incorporating spatial information. Integrates the underlying features extracted by traditional methods and deep learning features to train the classifier, considering the spatial relationships between the features. Guo et al. suggested a feature-fusion-based outdoor weather picture classification approach that includes sky, contrast, and saturation. According to [8], it has been found that deep features extracted by the dark channel and the AlexNet network exhibit a recognition accuracy rate exceeding 90%. The third type of method combines the underlying features and depth features of weather images, allowing for a more detailed expression of image information from various spatial perspectives. When comparing the first two types of methods, the addition of spatial information further enhances the accuracy of weather identification.

In order to address the issue of limited availability of public data sets for outdoor image weather recognition, as well as the lack of alignment between existing data sets and the specific application focus of this article, a new data set called the Multiclass Weather Image Blocks Dataset (MWIBD) was created. This data set includes various weather conditions such as fog, rain, snow, and sunny outdoor weather images. This study presents a novel approach to recognising

outdoor weather conditions using image blocking and feature fusion. It improves existing weather recognition methods and employs the migration VGG16 network model to extract deep features from weather images. These features are then combined with superficial features such as average gradient, contrast, saturation, and dark channel to train a SoftMax classifier. The classifier is capable of accurately identifying fog, rain, snow, and clear weather conditions.

The significance of the work is to increase the visibility of the weather image of the global spatial information system. Intelligent global transportation and video surveillance require this. The main goal is to develop and execute a strong weather detection algorithm that can quickly and reliably identify weather in various areas. Standard image processing retrieves surface characteristics, whereas pre-trained VGG16 models collect deep spatial features. The goal is to train a SoftMax classifier to distinguish clear, rain, snow, and fog using a spatially improved, multicategory weather picture block dataset using feature fusion. Finally, our surveillance film should be waterproof while providing accurate and reliable location data for these devices. Experimental validation shows the utility of the algorithm locally. The 99.26% weather recognition accuracy suggests that it might be effective in spatial information-based adaptive video picture clarification systems.

This study presents a novel perspective on the impact of inclement weather on geographical data used in intelligent transportation and video surveillance systems. The main accomplishments include a weather recognition system that takes into account geographical context, as well as a collection of weather images organized by place. The system can enhance its accuracy in distinguishing between four different weather conditions by utilising deep spatial information and typical image processing techniques. The accuracy of the algorithm, determined by testing, is 99.26%. The results demonstrate that the algorithm improves video clarity even in challenging circumstances. This approach improves the dependability of geographic data during adverse weather conditions.

The study proposes a complete method to mitigate the influence of unfavourable weather on GIS. Section 2 introduces a multicategory weather picture block dataset to facilitate model training and testing in the absence of acceptable datasets. A weather picture crashing dataset image. Geospatial feature extraction for weather photos is covered in Sect. 3. Deep features employ transfer learning, while shallow features use standard techniques. The unique strategy of merging shallow and deep information to improve the SoftMax classifier, which is critical for weather classification, is introduced in Sect. 4. The way in which the weather image recognition model distinguishes clear, snow, rain, and fog is explained in Sect. 5. In Sect. 6, Experimental Results and

Analysis, the system achieves 99.26% accuracy in weather identification. The algorithm's function in adaptive video picture clarification for spatial data integrity in bad weather is highlighted in Sect. 7, Conclusions, which highlights the important aspects and discusses future research.

2 Methods

2.1 Weather Image Chunking dataset

The currently available public weather image datasets, in terms of spatial coverage, are very limited. Chu et al. constructed the Image2 Weather dataset, which has 183,798 images. It includes photographs of clear, overcast, snow, rain, fog, and other 6 weather situations [9]. The Multiclass Weather Dataset (MWD) uses spatial data. There are 65,000 photos in 6 categories: sunny, overcast, rain, snow, fog, and thunder [10]. Although both public data sets have six categories of weather, it is observed that the presence of clouds or cloudy conditions does not significantly affect the recognition of video image targets. In the absence of any cover, such as trees or buildings, and with targets appearing blurred due to lightning weather conditions, it is worth noting that such weather occurrences are relatively rare. Therefore, the weather classified and identified in this article includes fog, rain, snow, and clear weather occurring in various locations.

This article selected some images that meet the needs of this article from the Image2 Weather dataset and MWD, and collected some public images to construct a weather image data set dataset. The sources of the data set are mentioned in Table 1.

The data set dataset includes four categories of images: fog, rain, snow, and clear, with 1,000 images in each category. Based on the purpose of the weather recognition model application in this article, the rainy and snowy images in the data set are those that are raining and snowing images. Since the data set is small and there are many identical targets and features in different types of weather image, such as roads, trees, vehicles, pedestrians, etc., it will seriously interfere with the recognition of weather in outdoor images. Therefore, this article chooses to crop the images. and flipping to filter out images with weather characteristics, thereby reducing interference and effectively improving weather recognition. This paper processes the data set dataset and constructs a multi-category weather image block data set

Table 1 Source distribution of weather image datasets

Weather Image Data Set	Foggy	Rainy	Snowy	Sunny
Image2 Weather Dataset	500	100	200	400
MWD	200	600	600	300
Public Image Dataset	300	300	200	300
	1000	1000	1000	1000

MWIBD. The processing process is shown in Fig. 1. The data set contains 4000 images, and each image is randomly cropped into 10 images with a size of 224 * 224.

First, each image in the data set is randomly cropped into 10 images of 224×224 size. There may be images without weather features in the cropped images, such as the close-up area of the foggy image. The fog is very light, approximately Sunny day images, so it is necessary to filter the images to remove images without weather features (fog, raindrops, snowflakes) or with unclear weather features, and then flip the remaining images left and right, finally forming a multi-category weather image block data set A sample of the MWIBD. MWIBD Fig. 2 shows data.

The MWIBD weather image block data set includes four categories of weather images: fog, rain, snow, and clear. Each category has 10,000 images and each image size is 224×224. 80% of the images in the MWIBD data set are used for training and 20% for training. Image used for testing.

2.2 Weather Image feature extraction

This paper extracts shallow features from weather images to express the spatial information of the image, and extracts deep features to express the abstract and intrinsic spatial information of weather images.

The three geographically separated modules that make up the weather recognition model are decision recognition, feature fusion, and feature extraction. Both shallow and deep feature extraction methods are included in the feature extraction module. From the provided spatial information, the shallow feature mainly derives the average gradient, contrast, saturation, and dark channel. The fully connected layer features FC1 and FC2 of VGG16GTL, which are situated inside the model's architecture, are the main sources of extraction for the deep feature.

2.2.1 Shallow-feature extraction

2.2.1.1 Average gradient The average gradient of a picture might indicate its clarity, according to this study [11]. Images with higher average gradients have more edge information and are sharper. The mean gradient of the grayscale image is as follows: Eq. (1) Shown:

$$G = \frac{\sum_{i=1}^w \sum_{j=1}^h \sqrt{\frac{(A_{ij} - A_{(i+1)j})^2 + (A_{ij} - A_{i(j+1)})^2}{2}}}{w * h} \quad (1)$$

Among them, A_{ij} indicates the image's pixel-point gray value (i, j) , w width and h height of the picture.

Fig. 1 Data Set Processing Process

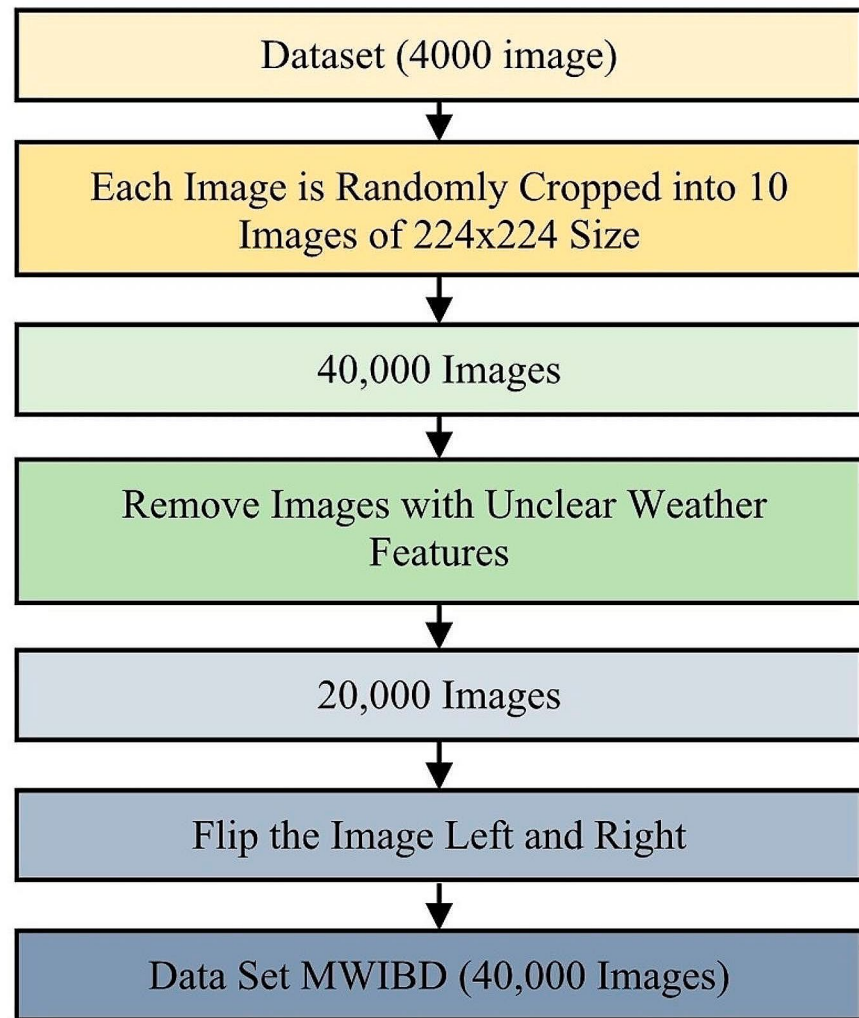


Fig. 2 Example of MWIBD Dataset

2.2.1.2 Contrast This study found that contrast describes changes in picture value in image space [11]. Degraded photos have less contrast than clean ones. Under different weather conditions, the contrast difference is large, so it can be used as a feature to distinguish the contrast of weather images. As shown in formula (2):

$$C = \sqrt{\frac{V}{\sqrt{\left[\sum_{k=1}^{255} (k - G)^4 * \frac{N_k}{w * h}\right] / V^2}}} \tag{2}$$

Among them, G is the average gradient of the image, V is picture standard deviation, N_k pixel count with gray value k , k represents input image gray value, $k \in [0.0.255]$. Among them, the standard of the image the difference V is as shown in Eq. (3):

$$V = \sqrt{\left[\sum_{i=1}^w \sum_{j=1}^h (A_{ij} - \bar{A}_{ij})^2\right] / (w * h)} \tag{3}$$

This includes, A_{ij} represents the mean value of A_{ij} .

2.2.1.3 Saturation This research indicated that saturation is not affected by lighting, so it can describe different images

under different lighting conditions [2]. The normalized saturation of each pixel of image I is as shown in Eq. (4):

$$S_{i,j} = \frac{S_{i,j} - \min(S_I)}{\max(S_I) - \min(S_I)} \tag{4}$$

Among them, $S_{i,j}$ represents the saturation of pixel point (i, j) , $\max(S_I)$ is the maximum saturation value of image I , and $\min(S_I)$ is the minimum saturation value of image I .

2.2.1.4 Dark passage The work suggested image dehazing using dark-channel prior theory. According to the prior hypothesis of the dark channel, most local portions of outdoor haze-free photographs include pixels with very low values in at least one color channel, approaching 0 [12]. Therefore, the dark channel represents hazy weather. Image I 's dark channel is Eq. (5):

$$I_d(x) = \min_{c \in \{R,G,B\}} (I_c(x)) \tag{5}$$

Among them, I_c represents the RGB channel of the image.

2.2.2 Deep feature extraction

2.2.2.1 Transfer learning Traditional machine learning and data mining work better when training and test sets are in the same feature space and equally distributed. Transfer learning is gaining popularity as each job requires a new data set, which is costly [13]. Transfer learning optimizes model training by transferring trained model parameters to a new model. Transfer learning improves training results, as most data and tasks are correlated [14]. This article fine-tunes transfer learning using the VGG16 network model.

2.2.2.2 VGG16 network model This research built VGG16 to solve ImageNet's 1000 picture categorization and placement categories [15]. Compared to AlexNet [16], VGG16 presents a smaller 3×3 and 1 network topology. The $\times 1$ convolution kernel and 2×2 pooling kernel can enhance network depth while reducing parameters. Figure 3 shows the structure of the VGG16 network model. VGG16 has 13 convolutional layers, 5 pooling layers, and 3 fully connected layers.

2.2.2.3 Deep feature extraction based on transfer learning The VGG16 network model has very good generaliza-

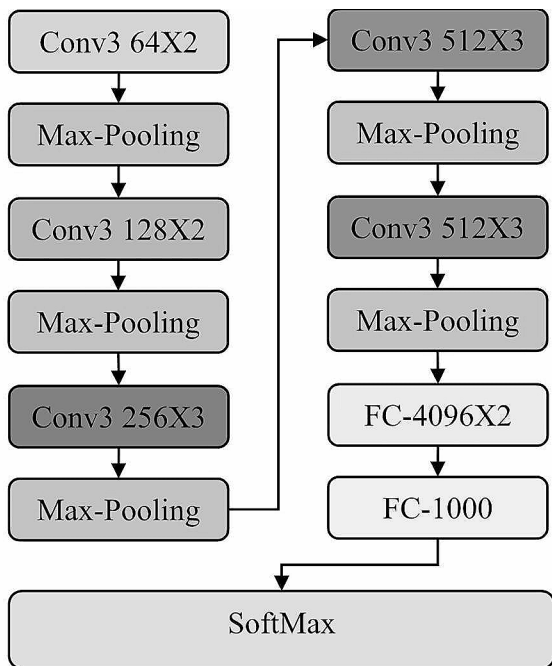


Fig. 3 VGG16 Structure Diagram

tion capabilities when migrating the network to new tasks. The model migration based on VGG16 is shown in Fig. 4.

The migration idea based on the VGG16 model is as follows: freeze 13 convolutional layers, remove all fully connected layers of the pre-trained model, and add 2 fully connected layers of your own design. Name the new two fully connected layers FC_1 and FC_2 respectively. Set the number of neurons output by FC_1 and FC_2 to 2048 and 1024 respectively.

The training parameters of the three fully connected layers FCG4096, FCG4096 and FCG1000 in the VGG16 network model are: $7 \times 7 \times 512 \times 4 = 96102760448$, $4096 \times 4096 = 6,777,216$, $4096 \times 1000 = 4,096,000$. And based on the training parameters of the two fully connected layers FC1 and FC2 of the VGG16 migration model, VGG16GTL are: $7 \times 7 \times 512 \times 2048 = 51,380,224$, $2048 \times 1024 = 2,097,152$. Fine-tuning VGG16 After that, the training parameters were greatly reduced, making training more efficient.

The optimizer used in this article to optimize the network model is Adam [17], which combines the advantages of two optimization algorithms, *AdaGrad* and *RMSProp*. The step size of Adam's calculation update is not directly

determined by the current gradient but by the first-order moment of the gradient. The estimation correction and the second-order gradient moment estimation correction are two parts of adaptive adjustment. Using Adam to update parameters is not affected by the scaling transformation of the gradient and can automatically adjust the learning rate. The relevant calculation formula for Adam is as follows.

The objective function $J(\theta)$ is differentiated with respect to θ to obtain the gradient g_t :

$$g_t = \nabla_{\theta} J(\theta_{t-1}) \quad (6)$$

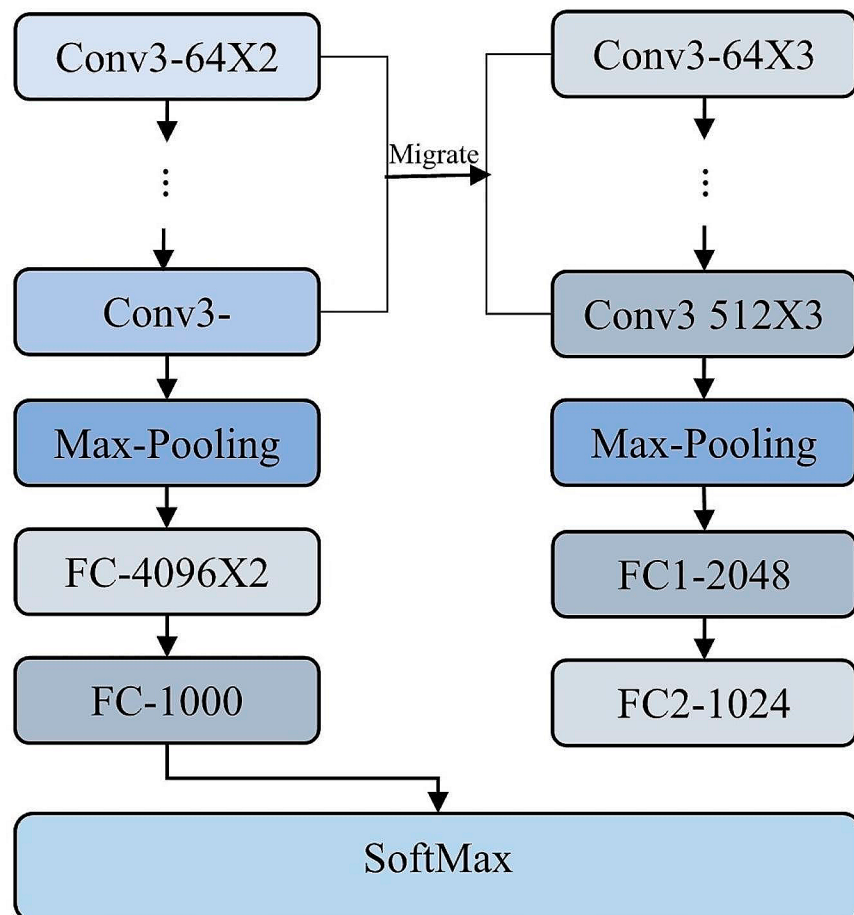
Among them, θ is the parameter of the network, which refers to the weight, deviation, or activation value; $J(\theta)$ refers to the objective function with parameter θ to be optimized.

The first-order moment estimates m_t of the gradient g_t is shown in Eq. (7):

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (7)$$

Among them, β_1 is the first-order moment attenuation coefficient.

Fig. 4 Migrate VGG16 Model



The second-order moment estimate v_t of gradient g_t is shown in Eq. (8):

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (8)$$

Among them, β_2 is the second-order moment attenuation coefficient.

Since the initial value of m_t is 0, it is biased towards 0 in the early stage of training, and m_t needs to be biased and corrected, as shown in Eq. (9):

$$\hat{m}_t = \frac{m_t}{(1 - \beta_1^t)} \quad (9)$$

Among them, β_1^t is the t^{th} power of β_1 .

Since the initial value of v_t is 0, it is biased towards 0 in the early stage of training, and v_t needs to be biased and corrected, as shown in Eq. (10):

$$\hat{v}_t = \frac{v_t}{(1 - \beta_2^t)} \quad (10)$$

Among them, β_2^t is the t^{th} power of β_2 .

The updated step size θ_t is shown in Eq. (11):

$$\theta_{t+1} = \theta_t - \alpha \frac{1}{\sqrt{\hat{v}_t + \epsilon}} \hat{m}_t \quad (11)$$

Among them, it was suggested through experiments to set $\epsilon = 10^{-8}$, and the learning rate α can be adjusted according to the specific situation [17]. This article sets it to 0.0001 on the basis of experiments.

The features of the fully connected layers FC_1, FC_2 in the transfer learning model based on VGG16 can better express the deep information of weather images, and the features of the fully connected layer are one-dimensional features, which are easy to integrate with traditional features. Therefore, this article extracts the FC_1 of VGG16-TL the features of the FC_2 layer is used as deep features of weather images.

2.3 Weather Image Feature Fusion

2.3.1 Feature Fusion Method

This article analyses in detail the fusion methods of different types of features and mentioned deep feature fusion methods, including additive fusion, maximum fusion, cascade fusion, etc. [18].

This paper adopts the idea of cascade fusion to fuse the shallow features and deep features of weather images. The feature fusion method is shown in Eq. (12):

$$F = [F_G, F_C, F_S, F_I, F_{FC_1}, F_{FC_2}] \quad (12)$$

Among them, F_G represents the average gradient feature, F_C represents the contrast feature, F_S represents the saturation feature, F_I represents the dark channel feature, F_{FC_1} and F_{FC_2} represent the deep features extracted from the fully connected layers FC_1, FC_2 of VGG16-TL respectively.

2.3.2 Weather Image Feature Fusion

The shallow features are to extract the average gradient, contrast, saturation, and dark channel histograms. Each feature has 256 dimensions, and the shallow features have a total of 1024 dimensions. The deep features are used to extract the features of the fully connected layers FC_1 and FC_2 of VGG16GTL, and the deep features have a total of 1024 dimensions. 3072 dimensions and then cascade and fuse all features to form 4096-dimensional weather image features.

This article uses TensorFlow in the feature fusion module. The `tf.concat()` function implements cascade fusion of average gradient, contrast, saturation, dark channel, FC_1 , and FC_2 .

2.4 Weather image recognition model

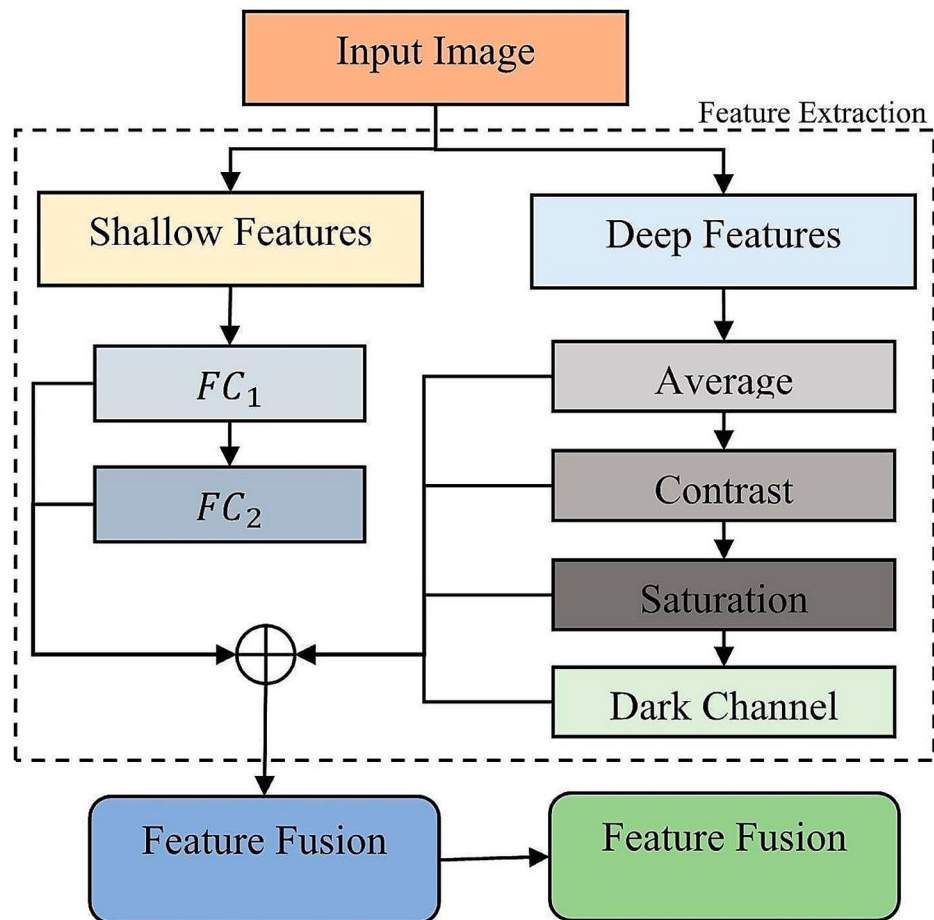
The outdoor image weather recognition model designed in this article based on image segmentation and feature fusion is shown in Fig. 5.

The weather recognition model consists of three modules that are spatially divided: feature extraction, feature fusion, and decision recognition. The feature extraction module consists of both deep feature extraction and shallow feature extraction techniques. The shallow feature mainly extracts the average gradient, contrast, saturation, and dark channel from the given spatial information. The deep feature mainly extracts the fully connected layer features FC_1 and FC_2 of VGG16GTL, which are located within the architecture of the model. The features can be found in Sect. 3, which provides specific details about the content of the extraction module.

The VGG16GTL transfer learning model, which is based on VGG16, has features of fully connected layers FC_1, FC_2 that can better describe the deep information of weather photos. Furthermore, the features of the completely connected layer are one-dimensional features that are simple to combine with traditional features. As a result, the FC_1 of VGG16-TL is extracted in this paper, and the properties of the FC_2 layer are used as deep features of the weather image.

The feature fusion module cascades and fuses the extracted deep features and the combined shallow features.

Fig. 5 Weather Recognition Model



Please, see Sect. 3 for the specific content of the feature fusion module. The decision recognition module is a feature fused through Softmax training, which can realize the recognition of four weather conditions: fog, rain, snow, and clear weather.

3 Results and discussion

3.1 Experimental environment

The experiments in this article were all conducted on an Intel(R) Core (TM) i7G9750H CPU @2.60 GHz, 8GB RAM, 64-bit Windows 10 system, NVIDIA GeForce GTX 1660 Ti, python 3. 6. Conducted in a TensorFlow2.2.0 GPU environment.

3.2 Evaluation index

The confusion matrix is an indicator to evaluate the effect of the model and can be used to evaluate the weather recognition model designed in this article [8]. The confusion matrix is listed in Table 2. In Table 2, T_P indicates that the image

Table 2 Confusion Matrix

		Actual Value	
		Positive Sample	Negative Sample
Predictive Value	Positive Sample	T_P	F_P
	Negative Sample	F_N	T_N

calibration is a positive sample and the classification result is also a positive sample. F_P means that the image is calibrated as a negative sample and the classification result is a positive sample; F_N means that the image is calibrated as a positive sample and the classification result is a negative sample; T_N means that the image is calibrated as a negative sample and the classification result is also a negative sample.

The weather recognition model of the article can be evaluated using a confusion matrix [8]. Table 2 shows the spatial confusion matrix. The geographical data in Table 2 and the classification result imply that T_P is a solid picture calibration and classification sample. F_P , F_N , and T_N indicate negative images with positive, negative, or negative classification results, respectively, in spatial calibration.

Precision(P): Represents the proportion of correctly classified samples among the samples identified as positive categories:

$$\text{Precision} = \frac{T_P}{T_P + F_P} \quad (13)$$

Recall (R): Represents the proportion of correct predictions among all positive category samples:

$$\text{Recall} = \frac{T_P}{T_P + F_N} \quad (14)$$

Accuracy: Indicate the proportion of correctly predicted samples to the total number of samples:

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + F_P + T_N + F_N} \quad (15)$$

3.3 Experiment 1: shallow feature experiment

This article designs Experiment 1 to compare the weather recognition of shallow features Local Binary Pattern (LBP), Histogram of Oriented Gradient (HOG), brightness, average gradient, contrast, saturation, and dark channel. Effect.

As a threshold, the local binary pattern (LBP) technique uses the gray value of the center pixel in a local picture region. Compare this threshold to nearby pixels. Its position is recorded as 1 if the neighbourhood pixel value is larger than the central location pixel value, suggesting a spatial link. If the neighbourhood pixel value is less than the central pixel value, the location is shown as 0, indicating a different spatial connection. From the binary number that expresses spatial information, a decimal number may be obtained. Image's center pixel LBP is stated in decimal form. The picture feature is a histogram of all local LBP values. HOG creates a histogram by computing gradient values in the immediate region of the image. Picture to create space. The experimental results of the shallow feature comparison are listed in Table 3.

Table 3 Shallow Feature Comparison

Shallow Features	Accuracy
LBH	36.75
HOG	42.50
Brightness	43.80
Saturation	51.83
Dark channel	55.08
Average Gradient	60.25
Contrast	62.66
Average Gradient + Contrast + Saturation + Dark Channel	74.66

As shown in Table 3, the impact of a shallow feature's recognition is poor, since it may be better at recognizing one sort of image than another. For example, the dark channel can effectively identify foggy weather images, but the classification effect on rainy and snowy days is poor; contrast can effectively distinguish foggy, sunny and rainy images, but it is difficult to distinguish between rainy and snowy images.

Although the classification effect of a single shallow feature is not good, the fusion of four shallow features can achieve a good classification effect. The fusion of the four shallow features of average gradient, contrast, saturation, and dark channel can achieve a recognition accuracy of 74.66%. Rate.

When comparing 7 shallow features, the accuracy of HOG, LBP, and brightness is relatively low. This article chooses to use four shallow features: average gradient, contrast, saturation, and dark channel.

3.4 Experiment 2: deep-feature experiment

Design experiment 2 compares the recognition effects of 4 network models *AlexNet* [16], *GoogLeNet* [19], ResNet50 [20], VGG16 [21] and the migration models of these 4 network models: Alex-Net-TL [22], *GoogLeNet - TL* [23], ResNet50-TL [24] and VGG16-TL [25].

The training parameter settings of all models are consistent, the batch size *batch_size* is set to 32, the network model optimizer is Adam, the learning rate is set to 0.0001, and the number of training iterations epoch is set to 100. 4 network models *AlexNet*, *GoogLeNet*, the change curves of the recognition accuracy of *ResNet50* and *VGG16* in different iteration cycles are shown in Fig. 6 (a) and 6 (b).

As can be seen in Fig. 6, the training loss of *GoogLeNet* is relatively the smallest among the four trained network models. The training loss after 20 epochs is basically less than 0.1, and the classification accuracy is high, up to 96.20%.

Comparing the four full training models, first of all, *AlexNet* uses large convolution kernels with sizes of 11×11 and 5×5 , while VGG16 uses smaller convolution kernels with sizes of 3×3 and 1×1 , and uses small convolution kernels. Replacing large convolution kernels with convolution kernels can deepen the network depth and reduce network parameters, so the effect of VGG16 is better than *Alex - Net*. Secondly, *AlexNet* and VGG16 both have convolutional, pooling, and fully linked layers. *ResNet50* uses the residual structure is replaced, while *GoogLeNet* uses the Inception Module structure and converts all connections to sparse connections [19]. Therefore, *GoogLeNet* not only effectively controls the parameters and calculation amount, but also has better recognition

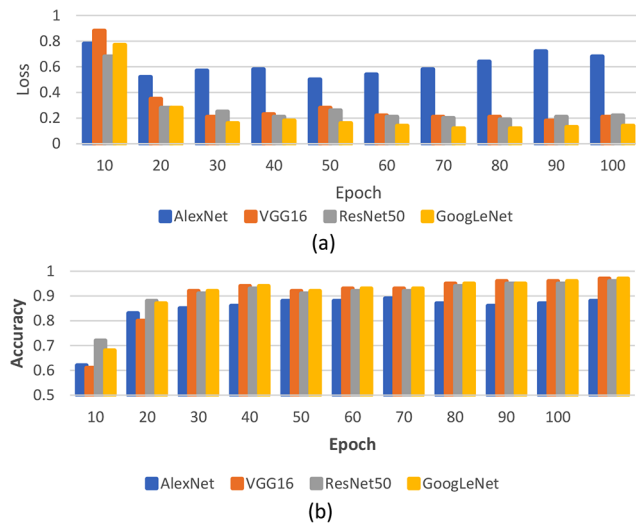


Fig. 6 Recognition loss and accuracy of 4 mitigation models in different iteration cycles. **a** Loss Changes Network Mode, **b** Accuracy Changes Network Mode

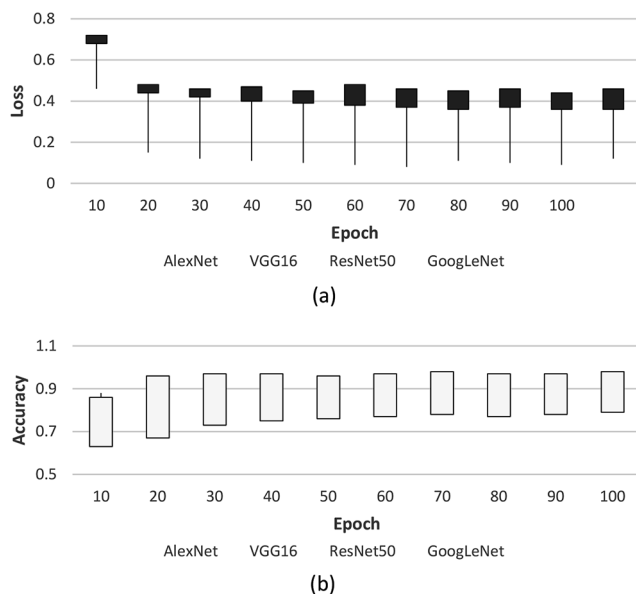


Fig. 7 Classification loss and accuracy of 4 migration models in different iteration cycles, **(a)** Loss Change Transfer Network Model **(b)** Accuracy Changes Transfer Network Model

results. After comparing the 4 full training models, compare the migration network models of these 4 models, namely *AlexNet – TL*, *GoogLeNet – TL*, *ResNet50 – TL*, *VGG16 – TL*. The classification accuracy change curves of these 4 migration models in different iteration cycles are shown in Fig. 7(a) and 7(b). As shown in Fig. 7, it can be seen that the convergence speed of the four models of the migration network is faster than that of the fully trained network model, but the training loss of *GoogLeNet – TL* and *ResNet50 – TL* is larger than that of the fully trained

Table 4 Recognition Accuracy of Network Model

Network Model	Accuracy
AlexNet – TL	79.93
AlexNet	85.77
GoogLeNet – TL	87.54
ResNet50 – TL	91.75
VGG16	94.85
ResNet50	95.26
GoogLeNet	96.22
VGG16 – TL	97.2

GoogLeNet and *Res – Net50*. When comparing the four migration network models, among them, the training loss of the *VGG16GTL* model is relatively minimal. After 10 epochs, the training loss is already lower than 0.1, and the classification accuracy is high, up to 97.29%.

Although among the four fully trained models, *GoogLeNet* has better recognition results, when performing transfer learning, *VGG16GTL* has better generalization ability and can achieve good transfer effects. Therefore, *VGG16GTL* extracts deep features from weather images CNN [26] preferred model weather images.

Comparing the training loss changes and the accuracy change curves of the full training model and the transfer model of a total of 8 models, the *VGG16 – TL* model has a relatively small loss and the fastest training convergence speed. The comparison of the specific recognition accuracy of the 4 network models and the 4 migration network models is listed in Table 4.

It can be seen in Table 4 that the recognition accuracy of the *VGG16GTL* model [27] is relatively the highest.

Taken together, the *VGG16-TL* model has smaller losses [28], the fastest training convergence speed, and the highest weather recognition accuracy. Therefore, this article chooses the *VGG16-TL* model to extract deep features of weather images [29].

3.5 Experiment 3: comparison of different methods

To verify the effectiveness and superiority of the proposed model, experiment 3 is designed to compare the proposed method with previous methods presented in articles [30], [31], and [32]. This article compares the recognition effects of four methods on four types of images: fog, rain, snow, and clear. The experimental results are listed in Table 5.

3.6 Discussion

As can be seen in Table 5, the overall recognition accuracy of the model in this document can reach 99.22%. The recognition effect on fog, rain, snow, and clear weather and the overall recognition accuracy are better than those of other

models. Among them, especially the recognition effect of rain and snow is greatly improved compared to other methods. The network structure used in Guo et al.'s method is *Alex-Net*. Through Experiment 2, it can also be seen that the weather recognition effect of *AlexNet* full training is poor. The method can extract deep features of images and express deep semantic information of images, but lacks the expression of shallow information of images. Therefore, although the recognition accuracy of Wang et al.'s method on the MWIBD data set in this paper has reached more than 90%, its recognition effect has yet to be improved. Improvement. This method can easily identify foggy days, but the recognition effect on rainy days and snowy days is poor. Goswami has proved through experiments that it is difficult to improve the weather recognition effect of the network model by increasing the number of convolution layers. Therefore, this paper integrates *VGG16-TL* extraction. Deep image features and shallow image features are used for weather recognition, which can achieve better weather recognition results. Taken together, the weather recognition model proposed in this article takes into account both shallow features and deep features of weather images, and achieves better recognition results in four types of weather: foggy, rainy, snowy, and sunny weather. However, the model in this article also has misidentifications. The model's recognition effect on rainy days and snowy days is not as good as foggy days and sunny days. This is because the two types of rainy and snowy days degradation are relatively similar and easily confused. The outdoor image weather recognition model designed in this article is intended to be applied in an adaptive video image sharpening processing system. When the current weather of the video image is recognized, the corresponding sharpening processing method is automatically called to improve the visibility of the degraded video image. If it is recognized that the current weather is raining, the video image rain removal method requires real-time processing to improve the visibility of video surveillance in a timely manner.

4 Conclusion

In this study, we created a weather image segmentation dataset that we refer to as MWIBD for outdoor weather image recognition. The dataset includes photographs of the weather from a wide variety of categories. In addition to this, it suggests a model for weather recognition that uses techniques such as image segmentation and feature fusion. This model can classify photographs of the weather by combining superficial and deep aspects of photographs through the use of spatial information. It does a good job of collecting both surface-level information, such as how the image

Table 5 Comparison of 4 Methods

Evaluation Index		Hierarchical Latent Features[30]	Cross-Layer Fusion[31]	Cascading Attention[32]	MWIBD
Fog	P	95.52	96.63	97.14	99.61
	R	95.91	96.42	98.17	99.54
Rain	P	89.33	90.56	96.14	99.05
	R	90.55	92.81	95.88	99.15
Snow	P	89.76	91.81	95.76	98.45
	R	89.96	92.32	96.64	99.38
Sunny	P	96.81	98.21	99.42	99.82
	R	94.82	95.72	97.78	99.18
Accuracy		92.84	94.35	97.22	99.22

appears, and the deeper level of semantic information, such as the meaning that lies behind the image. The model improves the accuracy of weather identification by using geographic data at every stage of the development process. The recommended weather identification model not only works well in a variety of geographical contexts, but also converges rapidly and has a low training loss. In the future, the method will try to be improved to be suitable for applications that include the management of spatial information, such as adaptive video image sharpening.

Acknowledgements Not applicable.

Funding Statement This work does not receive any kind of funding.

Data availability Data shall be made available on request.

Declarations

The authors have no conflicts of interest. No human or animal participation is involved in this research.

References

1. Lin, D., Lu, C., Huang, H., & Jia, J. (Sept. 2017). RSCM: Region Selection and Concurrency Model for Multi-class Weather Recognition. *IEEE Transactions on Image Processing*, 26(9), 4154–4167. <https://doi.org/10.1109/TIP.2017.2695883>.
2. Yu, T., Kuang, Q., Hu, J., Zheng, J., & Li, X. (2021). Global-Similarity Local-Saliency Network for Traffic Weather Recognition. *Ieee Access : Practical Innovations, Open Solutions*, 9, 4607–4615. <https://doi.org/10.1109/ACCESS.2020.3048116>.
3. Negru, M., Nedevschi, S., & Peter, R. I. (2015). Exponential contrast Restoration in Fog conditions for driving assistance. *IEEE Transactions on Intelligent Transportation Systems Aug*, 16(4), 2257–2268. <https://doi.org/10.1109/TITS.2015.2405013>.
4. Zheng, X., et al. (June 2019). Detecting comma-shaped clouds for severe Weather forecasting using shape and motion. *IEEE Transactions on Geoscience and Remote Sensing*, 57(6), 3788–3801. <https://doi.org/10.1109/TGRS.2018.2887206>.
5. Cheng, Y., Jia, Z., Lai, H., Yang, J., & Kasabov, N. K. (2020). A fast sand-dust image Enhancement Algorithm by Blue Channel Compensation and guided image Filtering. *Ieee Access : Practical Innovations, Open Solutions*, 8, 196690–196699. <https://doi.org/10.1109/ACCESS.2020.3034151>.

6. Xue, J., Zhang, H., Nishino, K., & Dana, K. J., Differential viewpoints for Ground Terrain Material Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 44(3): 1205–1218, 1 March 2022, <https://doi.org/10.1109/TPAMI.2020.3025121>.
7. Raghunandan, K. S., et al. (Sept. 2018). Riesz Fractional based model for enhancing license plate detection and recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(9), 2276–2288. <https://doi.org/10.1109/TCSVT.2017.2713806>.
8. Tourani, A., Shahbahrami, A., Soroori, S., Khazae, S., & Suen, C. Y. (2020). A Robust Deep Learning Approach for Automatic Iranian vehicle license plate detection and Recognition for Surveillance Systems. *Ieee Access : Practical Innovations, Open Solutions*, 8, 201317–201330. <https://doi.org/10.1109/ACCESS.2020.3035992>.
9. Zhao, C., Zhang, S., Luo, R., Feng, S., & Kuang, G. (2023). Scattering Features Spatial-Structural Association Network for Aircraft Recognition in SAR Images. in *IEEE Geoscience and Remote Sensing Letters*. 20: 1–5, Art no. 4006505, <https://doi.org/10.1109/LGRS.2023.3280442>.
10. Ratnasingam, S., & McGinnity, T. M. (Aug. 2012). Chromaticity Space for Illuminant Invariant Recognition. *IEEE Transactions on Image Processing*, 21(8), 3612–3623. <https://doi.org/10.1109/TIP.2012.2193135>.
11. Mahmoud, S. A., Afifi, M. S., & Green, R. J. (1988). Recognition and velocity computation of large moving objects in images. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 36(11), 1790–1791, Nov. <https://doi.org/10.1109/29.9020>.
12. Gao, F., Liu, Q., Sun, J., Hussain, A., & Zhou, H. (2019). Integrated GANs: Semi-supervised SAR Target Recognition. *Ieee Access : Practical Innovations, Open Solutions*, 7, 113999–114013. <https://doi.org/10.1109/ACCESS.2019.2935167>.
13. Xu, Y., Wen, J., Fei, L., & Zhang, Z. (2016). Review of video and image defogging algorithms and related studies on image restoration and enhancement. *Ieee Access: Practical Innovations, Open Solutions*. 4, 165–188. <https://doi.org/10.1109/ACCESS.2015.2511558>.
14. Wu, D., (2023). DCFusion: A Dual-Frequency Cross-Enhanced Fusion Network for Infrared and Visible Image Fusion. *IEEE Transactions on Instrumentation and Measurement*. 72, 1–15. Art no. 5011815. <https://doi.org/10.1109/TIM.2023.3267380>.
15. Cheng, Y., Jia, Z., Lai, H., Yang J. and Kasabov, N. K. (2020). Blue Channel and Fusion for Sandstorm Image Enhancement. *IEEE Access*. 8, 66931–66940. Art no. 5011815. <https://doi.org/10.1109/ACCESS.2020.2985869>.
16. B. Yu, Y. Chen, S. -Y. Cao, H. -L. Shen and J. Li. (2022). Three-Channel Infrared Imaging for Object Detection in Haze. *IEEE Transactions on Instrumentation and Measurement*. 8, 66931–66940. Art no. 5011815. <https://doi.org/10.1109/TIM.2022.3164062>.
17. Zhang, C., et al. (2019). Weather Visibility Prediction based on Multimodal Fusion. *Ieee Access : Practical Innovations, Open Solutions*. 7, 74776–74786. <https://doi.org/10.1109/ACCESS.2019.2920865>.
18. Wang, Y., Zhang, Z., Hao, W., & Song, C. (2021). Multi-domain image-to-image translation via a Unified Circular Framework. *IEEE Transactions on Image Processing*, 30, 670–684. <https://doi.org/10.1109/TIP.2020.3037528>.
19. Zou, Y., et al. (2020). A robust license plate Recognition Model based on Bi-LSTM. *Ieee Access : Practical Innovations, Open Solutions*, 8, 211630–211641. <https://doi.org/10.1109/ACCESS.2020.3040238>.
20. Panahi, R., & Gholampour, I. (April 2017). Accurate detection and Recognition of Dirty Vehicle Plate Numbers for high-speed applications. *IEEE Transactions on Intelligent Transportation Systems*, 18(4), 767–779. <https://doi.org/10.1109/TITS.2016.2586520>.
21. Huang, S. C., Chen, B. H., & Wang, W. J. (2014). Visibility Restoration of Single Hazy Images Captured in Real-World Weather Conditions. *IEEE Transactions on Circuits and Systems for Video Technology*. 24(10), 1814–1824, Oct. <https://doi.org/10.1109/TCSVT.2014.2317854>.
22. Fang, C., Song, Y., Guan, F., Liang, F., & Yang, L. (June 2023). Complex-valued deep neural network for Target Recognition of UAV SAR Imagery. *IEEE Journal on Miniaturization for Air and Space Systems*, 4(2), 175–185. <https://doi.org/10.1109/JMASS.2023.3247586>.
23. Chen, S., Shu, T., Zhao, H., & Tang, Y. Y. (2023). MASK-CNN-Transformer for real-time multi-label weather recognition. *Knowledge-Based Systems*, 278, 0950–7051.
24. Xie, K., Huang, L., Zhang, W., & Qin, Q. (2022). Lei Lyu,a CNN-based multi-task framework for weather recognition with multi-scale weather cues. *Expert Systems with Applications Volume*, 198, 0957–4174.
25. Wenchen Yang, Y., Zhao, Q., Li, F., & Zhu (2023). Yu Su,Multi visual feature fusion based fog visibility estimation for expressway surveillance using deep learning network, *Expert Systems with Applications*, Volume 234, 121151, ISSN 0957–4174.
26. S Divya Meena, Veeramachaneni Gayathri siva sameeraja, Nagineni Sai Lasya, Meda Sathvika, Veluru Harshitha, J Sheela. (2022). Hybrid Neural Network Architecture for Multi-Label Object Recognition using Feature Fusion. *Procedia Computer Science*. 215, 78–90. ISSN 1877–0509.
27. Kaihua, X., Kang, X., & Liu, H. (2023). Puhong Duan,MOFA: A novel dataset for multi-modal image Fusion Applications *Information Fusion*, 96, Pages 144–155, ISSN 1566–2535.
28. Tang, Q., & Liang, J. (2023). Fangqi Zhu, a comparative review on multi-modal sensors fusion based on deep learning *Signal Processing*. 213:109165, ISSN 0165–1684.
29. Chen, L., Zhan, W., Tian, W., He, Y., & Zou, Q. (2019). Deep Integration: A Multi-Label Architecture for Road Scene Recognition. *IEEE Transactions on Image Processing*. 28(10): 4883–4898, Oct. <https://doi.org/10.1109/TIP.2019.2913079>.
30. Wang, H., Zhang, Z., Hu, Z., & Dong, Q. (2022). SAR-to-optical image translation with hierarchical latent features. *IEEE Transactions on Geoscience and Remote Sensing*, 60(1-12), Art5233812. <https://doi.org/10.1109/TGRS.2022.3200996>.
31. Li, Qu, Z., & Wang, S. (2021). Ling Liu,a method of cross-layer fusion multi-object detection and recognition based on improved faster R-CNN model in complex traffic environment. *Pattern Recognition Letters*, 145, 127–134.
32. Wang, J., Huang, X., & Gai, S. (2019). Single image rain removal via cascading attention Aggregation Network on Challenging Weather conditions. *Ieee Access : Practical Innovations, Open Solutions*, 7, 178848–178861. <https://doi.org/10.1109/ACCESS.2019.2959041>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.