MFGB-S3W Denoiser: Multi-scale Fuzzy Granular-Ball and Sequential Three-Way Decision for Remote Sensing Image Denoising

Jingjing Qian[1] and Shuhua Su[1]

[1] School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Songjiang 201620, Shanghai, PR China
Qianjingjing716@163.com
Sushuhua913@163.com

Abstract. Existing denoising algorithms for remote sensing images (RSIs) often suffer from high noise sensitivity, insufficient retention of local details, and loss of information due to non-discriminatory denoising. Therefore, this paper proposes a denoising algorithm based on multi-scale fuzzy granular ball (MFGB) and sequential three-way decision (S3WD), denoted as the MFGB-S3W denoiser. First, a noise outlier assessment method is designed based on fuzzy rough set theory to quantify the sample noise outlier score to distinguish the noisy image from the normal image; second, the MFGB space is constructed to capture the global structural correlation and the local detail differences of the image through the multi-granularity feature fusion. Finally, combined with the S3WD mechanism, the image is segmented layer by layer to realize the precise localization of the noisy image. Experimental results demonstrate that the proposed model yields superior performance in terms of mAP_{50} , PSNR, SSIM, and ERGAS on the NWPU VHR-10 and RSOD remote sensing datasets, outperforming the Blind2Unblind method without MFGB and S3WD fusion, as well as the DnCNN and BM3D approaches. The proposed method effectively suppresses noise while preserving noise-free images, offering a promising solution for denoising complex RSIs.

Keywords: Fuzzy rough set , Sequential three-way decision, Granular ball, Image denoising

1 Introduction

As remote sensing technology rapidly advances, high-resolution RSIs play an Irreplaceable role in various fields such as environmental monitoring, land-use planning, and disaster assessment [1]. However, due to the combined effects of sensor noise, atmospheric interference, and transmission errors during the imaging process, the original RSIs are often severely affected by noise pollution. These noises lead to issues such as blurring of fine feature details and loss of edge information, which directly undermine subsequent recognition tasks and reduce the practical value of the

data [2]. Therefore, effectively removing noise while preserving key image features has become one of the core challenges within remote sensing informatics.

In recent years, deep learning (DL) has made significant progress, becoming a crucial methodology for RSI processing and recognition [3]. Research scholars are committed to employing DL techniques to enhance the performance of RSI recognition, and an array of enhanced algorithmic frameworks has been introduced to address the noise issue in RSIs. Many scholars improve the classical denoising algorithm BM3D [4]. For example, Chen et al. [5] introduce a localized patch correspondence framework leveraging edge-guided exploration and residual refinement mechanisms, improving BM3D to better handle the high similarity at the edges of RSIs. In addition, subsequent algorithms such as DGDN [6] and DLRP [7] also achieve excellent performance in the field of denoising. These algorithms typically denoise all images indiscriminately, which may result in over-denoising of clean images, resulting in degradation of primary data integrity and compromised accuracy in downstream recognition tasks.

Rough set theory, as an effective soft computing method, demonstrates superior performance in handling classification problems under uncertainty. Since the introduction of the 3WD method by Yao [8], which is grounded in rough set theory, it has been widely adopted by researchers in the domains of classification and anomaly detection. Yuan et al. [9] develop an incremental learning paradigm grounded in progressively fuzzified three-way principles. Luo et al. [10] propose a 3WD model basis integrated learning. Ying et al. [11] develop the NCDM denoising module by integrating the S3WD framework. However, these traditional methods often struggle to capture multi-scale information when processing complex data. To address this issue, many researchers have incorporated granular computing to enhance data processing efficiency. Xia et al. [12] propose the theory of granular balls based on granular computing, which specifies both granular ball labeling and inter-ball distance metrics, and further derives the foundational models of GBSVM and GBkNN. Building upon these developments, Gao et al. [13] designed an outlier detection methodology operating across scales via fuzzy rough approximations to accurately identify various types of outliers. Zhang et al. [14] investigate three-way neighborhood feature regions and corresponding fusion measures to enhance outlier detection. However, the aforementioned methods rely solely on a single application of 3WD for detection, overlooking dynamic decision-making scenarios characterized by uncertai-nty and incomplete information.

Therefore, to overcome inherent constraints in current denoising methodologies and simultaneously enhance the performance of RSI recognition, this paper proposes a novel denoising framework by integrating MFGB and S3WD. The resulting MFGB-S3W denoiser is capable of accurately identifying noisy images, thereby improving the overall denoising effect. The main contributions of this work are summarized as follows:

(1) To evaluate the noise level of samples, a fuzzy rough set-based method is developed to accurately identify noise outliers.

(2) To better capture both local details and global structures of RSIs, a MFGB space is constructed.

(3) By integrating MFGB and S3WD, and incorporating a self-supervised learning strategy, an efficient denoising algorithm is proposed that not only accurately identifies noisy images but also avoids over-denoising in noise-free images.

The subsequent sections develop as follows. Section 2 introduces the theoretical background. Section 3 elaborates on the proposed MFGB-S3W denoiser. Section 4 demonstrates the effectiveness of the MFGB-S3W denoiser through comprehensive experiments. Section 5 provides the conclusion of the study and outlines potential directions for future research.

2 Preliminaries

2.1 Fuzzy Rough Set

If there are no decision attributes in A, then the data is called a fuzzy information system and is defined as $FIS = (U, A, V, f)$, where $V = \bigcup_{a \in A} V_a$. On each attribute $a \in A$, the mapping function f relates each sample $x \in U$ to the value of attribute a, which can be expressed as $f(x, a) \in V_a$.

Definition 1. [15] The definition of fuzzy relation R on U is as $R : U \times U \mapsto [0,1]$, and it satisfies the following properties:

$$\begin{aligned} &R(x,x) = 1; \\ &R(x,y) = R(y,x); \\ &R(x,z) \geq \vee_{y \in U} (R(x,y) \wedge R(y,z)), \end{aligned} \tag{1}$$

where $R(x, y)$ represents the fuzzy similarity between samples x and y under R.

Definition 2. [13] Let attribute subset $B \subseteq A$ generate fuzzy relation R_B through induction. It can be calculated by $R_B(x_i, x_j) = \wedge_{a \in B} R_a(x_i, x_j)$,where $R_a(x_i, x_j)$ is defined as follows:

$$R_a(x_i, x_j) = \begin{cases} 1 - \left| f_a(x_i) - f_a(x_j) \right|, & \left| f_a(x_i) - f_a(x_j) \right| \leq \varepsilon_a, \\ 0, & \left| f_a(x_i) - f_a(x_j) \right| > \varepsilon_a, \end{cases} \tag{2}$$

where ε_a is a parameter threshold, computed from $\varepsilon_a = \frac{std(a)}{\delta}$, while $std(a)$ signifies the standard deviation of attribute a, and δ represents an adjustable parameter.

Definition 3. [13] As for all sets of attribute $B \subseteq A$, its fuzzy relation with density can be defined as:

$$U / \widetilde{R}_B = \{ [\widetilde{x}_1]_B, [\widetilde{x}_2]_B, ..., [\widetilde{x}_n]_B \} , \tag{3}$$

where $[\widetilde{x}_i]_B = (\widetilde{R}_B(x_i, x_1), \widetilde{R}_B(x_i, x_2), ..., \widetilde{R}_B(x_i, x_n))$ is denoted as the fuzzy granule with density of sample x_i ,and $\widetilde{R}_B(x_i, x_j) = \wedge_{a \in B} \widetilde{R}_a(x_i, x_j)$.

The cardinality of $[\widetilde{x}_i]_B$ is calculated as $\left|[\widetilde{x}_i]_B\right| = \sum_{j=1}^{n} \widetilde{R}_B(x_i, x_j)$, $1 \le \left|[\widetilde{x}_i]_B\right| \le n$.

Definition 4. [13] For any sample $x \in U$, with respect to attribute $a \in A$, the fuzzy granule density is given by:

$$Den_a(x) = \frac{\left|[x]_a\right|}{|U|}, \tag{4}$$

where $\left|[x]_a\right|$ represents the cardinality of the fuzzy granule $[x]_a$.

Definition 5. [13] To any two samples x_i and x_j , the relative fuzzy granular density of them with respect to attribute $a \in A$ is defined as :

$$\mathrm{Re}l\_Den_a(x_i, x_j) = \exp\{-\lambda \left\| Den_a(x_i) - Den_a(x_j) \right\|_2^2\}, \tag{5}$$

where λ is a weighted parameter, and $\|\cdot\|_2$ is 2-norm.

2.2 S3WD

S3WD integrates multi-granularity structures to enable dynamic decision-making. Its core idea is to construct a hierarchical granularity framework that iteratively makes decisions regarding objects inside the boundary region, progressing from coarse to fine levels.

Definition 6. [16] Let $FIS = (U, A, V, f)$ be a fuzzy information system and $A = \{A_1, A_2, ..., A_m\}$ be an attribute subset family with $A_1 \subseteq A_2 \subseteq ... \subseteq A_m \subseteq A$. Then STWD can be expressed as:

$$GS_i = (U_i, A_i \cup D, V_i, f_i), \tag{6}$$

$$GS = (GS_1, GS_2, ..., GS_q), \tag{7}$$

where $i = 1,2,...,m$, GS_i represents the *i*-th granularity space of *GS*,and *GS* expresses the multi-granularity space.

Let $X_i \subseteq U_i$ and the threshold of *i*-th layer be (α_i, β_i) , the three regions of the *i*-th layer can be expressed as:

$$\begin{aligned} POS_{(\alpha_i,\beta_i)}(X_i) &= \{x \in U_i \,|\, v_i(x) > \alpha_i\}; \\ BND_{(\alpha_i,\beta_i)}(X_i) &= \{x \in U_i \,|\, \beta_i \le v_i(x) \le \alpha_i\}; \\ NEG_{(\alpha_i,\beta_i)}(X_i) &= \{x \in U_i \,|\, v_i(x) < \beta_i\}, \end{aligned} \tag{8}$$

where U_i indicates the *i*-th layer domain, and $v_i(x)$ is the *i*-th evaluation function.

3 The Proposed Method

In real-world RSI recognition tasks, noise degrades image clarity and negatively impacts recognition performance. To enhance model performance, it is common practice to apply denoising to input images based on their noise levels before performing the recognition task. Such unmediated processing may excessively alter noise-free images, reducing downstream detection efficacy. In order to solve this problem, this paper proposes the MFGB-S3W Denoiser, which identifies and selectively denoises noisy images to obtain high-quality RSIs, thereby enhancing the training effectiveness of subsequent object detection tasks.

Specifically, a noise outlier score is first designed based on fuzzy rough set theory to quantify image noise. Then, MFGB space is constructed to comprehensively detect noise at different scale. Finally, the S3WD framework is constructed based on the previous two parts to distinguish the noisy images, and the Blind2Unblind self-supervised denoising algorithm is utilized to achieve efficient denoising.

The working procedure of the MFGB-S3W denoiser model proposed in this paper and the subsequent object detection task completion process are shown in Fig. 1. The pseudocode for the MFGB-S3W algorithm is shown in Fig. 2.

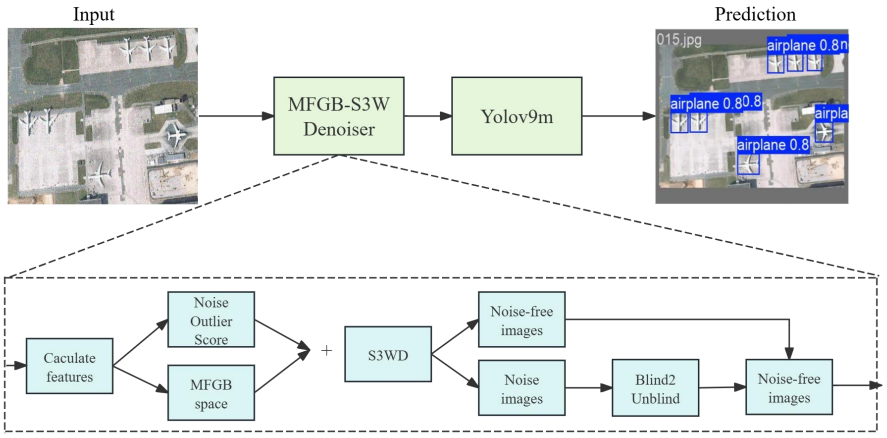

Fig. 1. Complete workflow diagram of the model in this paper

Algorithm MFGB-S3W Denoiser

Input: All images $U=(x_1,x_2,...,x_n)$, similarity threshold δ,weighted parameter λ, and the noise ratio estimation r_i.

Output: Noise-free images

Initialization: $GBS=\{GB_1,GB_2,...,GB_n\}$

Calculate: MEA, STD, SOB, LAP, ASM, ENT, HOM, DIS for all images to constitute a fuzzy information system $FIS=(U,A,V,f)$.

Calculate: The fuzzy granule density $Den_a(x_i)$, relative fuzzy granule density $\mathrm{Re}l\_Den_a(x_i,x_j)$ and fuzzy similarity degree with density $\widetilde{R}_a(x_i,x_j)$ by Eqs. (3)-(8).

for $i=1\rightarrow m$ **do**

  calculate $NS_{A_i}(x)$ by Eqs. (9)-(15).

end for

for $i=1\rightarrow m$ **do**

  for each pair (GB_o,GB_j) in GBS

    if $dis_{A_i}(GB_o,GB_j)\le\varepsilon_{A_i}$

      merge GB_o and GB_j

    else continue

  end for

end for

return MFGB space

while $BND_{(\alpha_i,\beta_i)}(X_i)=\phi$

for $i=1\rightarrow m$ **do**

  for $e=1\rightarrow E$

    calculate $P_i^e(x)$, v_i^e, H_i^e by Eqs. (19)-(20)

  end for

  calculate $\mu_i(x)$, $P_i(x)$ by Eqs. (21)-(22)

return α_i and β_i

determine the three-way regions by Eq. (24)

end for

end while

return $POS(X)=POS_{(\alpha_1,\beta_1)}(X_1)\cup...\cup POS_{(\alpha_z,\beta_z)}(X_z)$

Input the noisy images $POS(X)$ into Blind2Unblind algorithm

The denoised images $POS(\widetilde{X})$ are obtained

Return Noise-free images $\widetilde{X}=POS(\widetilde{X})\cup NEG(X)$

Output Noise-free images

Fig. 2. Pseudocode for MFGB-S3W Denoiser.

3.1 Noise Outlier Score

In the subsection, we proposes a method for calculating the noise outlier score. This score is constructed based on both the fuzzy granularity density of the attributes and their significance. A larger score indicates a higher likelihood that the sample is a noisy image. The whole calculation process is described as follows.

First, all images $U=(x_1,x_2,...,x_n)$ are input, and their attribute values including means (MEA), standard deviation (STD), Sobel (SOB), Laplacian (LAP), angular second moment (ASM), entropy (ENT), homogeneous (HOM), and dissimilarity (DIS), are computed [18].

Second, for all attributes $a\in A$, the fuzzy granule density $Den_a(x_i)$, relative fuzzy granule density $\mathrm{Re}l\_Den_a(x_i,x_j)$, and fuzzy similarity degree with density

$\widetilde{R}_a(x_i, x_j)$ are calculated according to Equation (2)-(5). In addition, for an attribute set $B \subseteq A$, the cardinality of the fuzzy granule with density $\left|[\widetilde{x}_i]_B\right|$ is calculated according to Equation (3), and thereby calculate the significance of *B* as follows [13]:

$$Sig(B) = -\log \sum_{x \in U} \frac{\left|[\widetilde{x}]_B\right|}{|U|}. \tag{9}$$

Third, all attributes $a \in A$ and subsets of attributes $A_i \subseteq A$ are ranked for significance. The specific sorting rules are as follows [13]:

$$AQ = \left\langle a_1^{'}, a_2^{'}, ..., a_m^{'} \right\rangle, \tag{10}$$

where $Sig(\{a_i^{'}\}) \geq Sig(\{a_{i+1}^{'}\})$. From this, it can be obtained.

$$ASQ = \left\langle A_1, A_2, ..., A_m \right\rangle, \tag{11}$$

where A_i is an attribute subset form of the top *i* attributes in *AQ*.

Finally, the noise outlier score (NS) for each image under all attribute subsets can be calculated as described in the following definition.

Definition 7. The NS of sample $x \in U$ with respect to each attribute subset $A_i \subseteq A$ are obtained as follows:

$$NS_{A_i}(x) = 1 - Sig(A_i) \frac{\left|[\widetilde{x}]_{A_i}\right|}{|U|}. \tag{12}$$

In this subsection, the calculation of significance for all attribute subsets $O(|A||U|^2)$ is an upper bound on time complexity. For each sample, the time to compute its noise outlier score is $O(|A|)$. As a result, the time complexity for this part can be expressed as $O(|A||U|^2)$.

3.2 MFGB Space

Granular ball computing, built upon information granularity, uses granular balls to model data samples for clustering and classification purposes.

Definition 8. [17] Let *GB* be a granular ball containing similar samples from *U* and $GBS = \{GB_1, GB_2, ..., GB_n\}$ be the set of granular balls. For any granular ball $GB_i \in GBS$, the center and radius of GB_i are defined as:

$$c_i = \frac{1}{|GB_i|}\sum_{x\in GB_i} x,$$
$$r_i = \max_{x\in GB_i} \|x - c_i\|_2. \tag{13}$$

Meanwhile, multi-scale learning serves as an effective strategy to enhance model performance. To more accurately identify noisy images, we construct granular-ball clustering structures at different scales based on granular ball computing theory, thereby forming a MFGB space. The detailed construction process is described as follows.

To begin with, initialize the set of granular-balls $GBS = \{GB_1, GB_2, ..., GB_n\}$ by considering each sample as an independent granular-ball.

Then, calculate the fuzzy similar degree between all granular balls two by two with respect to attribute $a \in A$. The specific formula is as follows:

$$R_a(GB_o, GB_j) = \begin{cases} 1 - dis_a(GB_o, GB_j), & dis_a(GB_o, GB_j) \le \varepsilon_a, \\ 0, & otherwise, \end{cases} \tag{14}$$

where

$$GB_o, GB_j \in GBS,\ dis_a(GB_o, GB_j) = \max(\left|f_a(c_o) - f_a(c_j)\right| - \left|r_o^{\frac{1}{|A|}} + r_j^{\frac{1}{|A|}}\right|, 0).$$

According to Equation (14), the fuzzy similarity relations of granular-balls under individual attributes can be aggregated into the fuzzy similarity relation under an attribute subset, as detailed below.

Definition 9. Let $A_i \subseteq A$ and $a_q \in A_i$, the fuzzy similarity relation under A_i is defined as:

$$R_{A_i}(GB_o, GB_j) = \frac{\sum_{q=1}^{k} R_{a_q}(GB_o, GB_j)}{k}, \tag{15}$$

where k is the number of attributes in A_i.

When $dis_{A_i}(GB_o, GB_j) \le \varepsilon_{A_i}$ where $o \ne j$, these two granular balls aggregate into a single granular ball. Thus we can get the granular ball at the i-th scale.

Ultimately, we iterate through all the subsets of the attributes $A_i \in A$ to obtain a collection of E scales granular balls, which ultimately forms a MFGB space.

In this subsection, the computational cost for granular ball initialization is $O(n)$, which in turn computes the fuzzy similarity of a single attribute, at which point the number of granular ball pairs is about n^2, and the total time complexity amounts to $O(|A|n^2)$. Next, the subset similarity of the attributes is aggregated, at which point the time complexity is $O(|A|n^2)$. The process of merging the granular ball is at most

$O(n^2)$ times and at least $O(n)$ times, so the time complexity is at most $O(n^2)$. In summary, the total computational complexity associated with this subsection is $O(|A|n^2)$.

3.3 MFGB-S3W Denoiser

The multi-scale space provides more comprehensive information about the data, enabling more effective detection of both global and local noise. Therefore, based on the previous two subsections, we propose a denoising method that fuses S3WD with MFGB. The denoising module operates according to the following detailed procedure.

First, the noise outlier scores of the samples at different scales are computed and converted to noise outlier probabilities, the conversion process is defined as follows.

Definition 10. For any sample $x \in U_i$, the possibility of noise outlier $P_i^e(x)$ in the e-th scale is defined as follows:

$$P_i^e(x) = \begin{cases} \frac{NS_i^e(x) - NS_i^{e_{t_i}}}{2(\max(NS_i^e) - NS_i^{e_{t_i}})} + \frac{1}{2}, & NS_i^e(x) > NS_i^{e_{t_i}}, \\ \frac{NS_i^e(x) + \min(NS_i^e)}{2(NS_i^{e_{t_i+1}} - \min(NS_i^e))}, & otherwise, \end{cases} \tag{16}$$

where $NS_i^e(x)$ denotes the noise outlier score of sample x under attribute subset A_i at the e-th level scale, $NS_i^{e_{t_i}}$ denotes the t_i -th largest value in the outlier score vector $NS_{A_i}^e$, and t_i is the number of noisy images obtained from the noise ratio estimation r_i ,which is computed by the product of the ratio r_i and the sample size $|U_i|$.

Second, in order to better integrate the effects of the different scales, the weights of each sample at each scale will be calculated. Under the attribute subset A_i, the weight of the e-th scale and the weight of the sample x are calculated as follows [13]:

$$v_i^e = 1 - \frac{\sum_{x \in U} H_i^e(x)}{U_i}, \tag{17}$$

$$H_i^e(x) = -P_i^e(x) \log P_i^e(x) - (1 - P_i^e(x)) \log(1 - P_i^e(x)), \tag{18}$$

where $H_i^e(x)$ is the average information entropy of x.

$$\mu_i(x) = 1 - \frac{\sum_{e=1}^{E} v_i^e \cdot H_i^e(x)}{E}, \tag{19}$$

where E is the number of scales.

When the noisy image is not clearly distinguishable from the normal image, the probability of noise anomaly of the sample is often unstable close to 0.5. Under these

circumstances, the average information entropy of the sample is close to 1, and the weight is close to 0.

Third, compute the threshold pairs (α_i, β_i) and the weighted noise outlier probability $P_i(x)$ for all samples within attribute subset A_i, as follows [13]:

$$P_i(x) = \frac{\sum_{e=1}^{E} v_i^e \cdot P_i^e(x)}{\sum_{e=1}^{E} v_i^e}, \tag{20}$$

$$\begin{aligned} \alpha_i &= OP(\lceil |U|(1-r+\Delta r) \rceil), \\ \beta_i &= OP(\lceil |U|(1-r-\Delta(1-r)) \rceil), \end{aligned} \tag{21}$$

where $OP(s)$ is the s-th probability of the noise outlier, and Δ is a parameter responsible for specifying the field size.

Fourth, according to the equation (8), all images will be divided into three regions $POS_{(\alpha_i,\beta_i)}(X)$, $BND_{(\alpha_i,\beta_i)}(X)$ and $NEG_{(\alpha_i,\beta_i)}(X)$ to get the decision result of the i-th layer. The positive region is the noisy outlier probability greater than α_i is regarded as noise images, and the probability lower than β_i is the noise-free images. Samples with noise outlier probability between α_i and β_i are considered uncertain samples and are put into the boundary region, requiring further identification.

Thus the samples in $BND_{(\alpha_i,\beta_i)}(X_i)$ are used as the sample set U_{i+1} in the i+1 level, where $U_{i+1} \subseteq U_i$. The final iteration ends when the bounding region is empty.

Fifth, aggregate the positive region samples obtained from each layer to obtain the set of all noisy images, which is defined as follows:

Definition 11. The set of all noisy images POS (X) is obtained based on the decision making results of each layer.

$$POS(X) = POS_{(\alpha_1,\beta_1)}(X_1) \cup ... \cup POS_{(\alpha_z,\beta_z)}(X_z), \tag{22}$$

where z is the number of layers on which the iteration stops.

Ultimately, all images classified as noise are denoised using Blind2Unblind [19]. Noise-free images are not processed to avoid excessive denoising.

In here we consider the maximum time complexity of the S3WD process, i.e., the subset of attributes all iterated through, which in reality will be less than this computed time complexity. In the computation of Equation (16)-(20), the time complexity is $O(|U_i||A_i|E)$ for all of them. The time complexity in determining the confidence threshold is $O(|A||U_i|\log|U_i|)$. Finally, when performing sequential three-way decision iterations, the maximum time complexity is $O(|A||U|\log|U|)$.

Combining the previous two subsections, the cumulative time cost of the algorithm in this paper is $O(|A||U|^2)$.

4 Experiment

4.1 Datasets

This paper selects the NWPU VHR-10 [20] and RSOD [21] high-resolution RSI benchmark datasets for validation. Comprising 650 target-containing images and 150 without targets, the NWPU VHR-10 dataset provides both positive and negative samples for evaluation. Each positive image includes at least one target instance, with a total of 3,651 annotated instances spanning 10 categories, such as aircraft, ships, and baseball fields. Additionally, the RSOD dataset consists of 976 aerial and satellite images focusing on four types of infrastructure targets. It provides pixel-level annotations, offering a valuable benchmark for object detection research.

Examples cover 10 different object classes, such as airplanes, ships, and baseball-diamond. We randomly divide the data into 80% for training and 20% for testing, and further extract 20% of the training set as the validation set. To simulate a noisy environment, the noise ratio is set to 30%, i.e., for each category of image 30% to add random Gaussian noise ($\delta = 25$). The effect of adding noise to the image is shown in Fig. 3.

Fig. 3. Comparison of noisy images and noise-free images.

4.2 Experimental Setup

Experimental Details. This experiment was conducted on an Ubuntu 18.04 LTS system using version name CUDA11.1, PyTorch 1.8.1 and four NVIDIA RTX 3090 GPUs with 24G of graphics memory on a single card. Before denoising the test image using the Blind2Unblind algorithm, pre-training was performed using the ImageNet validation set into noise, with the batch size set to 8, and an initial learning rate of 0.001. During the training of the YOLOv9m [22] object detection algorithm, the initial learning rate was also set to 0.001. A learning rate decay strategy was applied,

where the learning rate was reduced by a factor of 10 after 100 epochs. The training lasted for a total of 150 epochs with a batch size of 16.

Evaluation Metrics. In this paper, we mainly use the following metrics to measure the experimental effects, including mAP_{50} , Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) and Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS). The corresponding formula is as follows.

mAP_{50} is the average precision AP for all categories at an IOU threshold of 0.5.

$$AP = \int_0^1 P(R)dR\,, \tag{23}$$

where $P = \frac{TP}{TP+FP}$, $R = \frac{TP}{TP+FN}$.

$$PSNR = 10 \cdot \log_{10} \frac{(s_{l\max})^2}{MSE(s_l, \hat{s}_l)}\,, \tag{24}$$

where s_l is noise-free image, $\hat{s}_l$ is denoised image, and $s_{l_{\max}}$ is the maximum value of s_l .

$$SSIM = \frac{(2\mu_{\hat{s}_l}\mu_{s_l} + c_1)(2\sigma_{\hat{s}_l s_l} + c_2)}{(\mu_{\hat{s}_l}^2 + \mu_{s_l}^2 + c_1)(\sigma_{\hat{s}_l}^2 + \sigma_{s_l}^2 + c_2)}, \tag{25}$$

where μ_{s_l} and $\mu_{\hat{s}_l}$ are the mean of s_l and $\hat{s}_l$ respectively. σ_{s_l} and $\sigma_{\hat{s}_l}$ is the standard deviation of s_l and $\hat{s}_l$. c_1 and c_2 are constants.

$$ERGAS = 100 \times \frac{h}{l} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{RMSE_i}{\mu_i}\right)^2}\,, \tag{26}$$

where N is the number of bands, $RMSE_i$ is the RMSE of the i-th band, μ_i is the mean value of the i-th band, and $\frac{h}{l} = 1$.

4.3 Comparative Analysis

In the experimental evaluation, a comparison is carried out among BM3D [4], DnCNN [23], Blind2Unblind [19] , and the proposed approach using the NWPU VHR-10 and RSOD datasets. The detailed results are presented in **Table 1**. As summarized in **Table 1**, our model demonstrates superior performance over other denoising algorithms on both the NWPU VHR-10 and RSOD datasets. On the NWPU VHR-10 dataset, our model achieved an mAP_{50} of 86.4%, representing a 0.6% improvement over Blind2Unblind, which does not incorporate the MFGB-S3W framework. In addition, it attained a PSNR of 32.2 dB, a SSIM of 87.7%, and an ERGAS value of 10.6. On the RSOD dataset, our model further demonstrated its

robustness by achieving an mAP_{50} of 93.6%, a PSNR of 33.5 dB, and an SSIM of 89.8%. Meanwhile, the ERGAS value decreased to 6.6, indicating improved radiometric fidelity. Overall, the proposed denoising method not only enhanced image quality but also significantly improved the performance of object detection tasks.

Table 1. Performance comparison of various denoising algorithms on different datasets

Denoising algorithm	NWPU VHR-10 Dataset				RSOD Dataset			
	mAP_{50}	PSNR	SSIM(%)	ERGAS	mAP_{50}	PSNR	SSIM(%)	ERGAS
BM3D	83.4	29.9	80.7	11.6	86.3	30.6	83.4	7.5
DnCNN	81.1	30.1	79.3	11.4	89.6	31.3	83.9	6.9
Blind2Unblind	85.8	31.4	87.5	10.7	92.9	33.4	87.1	6.7
Our model	**86.4**	**32.2**	**87.7**	**10.6**	**93.6**	**33.5**	**89.8**	**6.6**

Fig. 4. shows the before and after comparisons of the noiseless image processed by excessive denoising, and it can be found that the image after excessive denoising loses a lot of original details. Meanwhile in Fig. 5, the four states of an image with noise-free, noisy, denoising by DnCNN, BM3D and our model are shown from left to right, which visually demonstrates the degradation caused by noise and highlights the varying effectiveness of different denoising algorithms.

Fig. 4. Noise-free image after over denoising.

Fig. 5. Visual comparison of noise-free images, noisy images, and denoised images.

5 Conclusion

In this paper, the proposed MFGB-S3W denoiser is a denoising method integrating MFGB and S3WD, which can accurately identify noisy regions and realize adaptive denoising. Experimental results show that MFGB-S3W outperforms existing algorith-

ms in mAP_{50} , PSNR SSIM and ERGAS metrics on RSI datasets. The proposed method effectively avoids excessive smoothing in noisy regions. In addition, the model can be combined with other denoising algorithms to further explore its practical value and application potential.

References

1. Mena F., Arenas D., Nuske M., Dengel A.: Common practices and taxonomy in deep multiview fusion for remote sensing applications. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 17, 4797-4818 (2024). doi:10.1109/JSTARS.2024.3361556
2. Feng W., Long Y., Wang S., Quan Y.: A review of addressing class noise problems of remote sensing classification. Journal of Systems Engineering and Electronics 34(1), 36-46 (2023). doi: 10.23919/JSEE.2023.000034
3. Cheng X., Sun Y., Zhang W., Wang Y., Cao X., Wang Y.: Application of deep learning in multitemporal remote sensing image classification. Remote Sensing 1(15), 385 (2023). doi: 10.3390/rs15153859
4. Danielyan A., Katkovnik V., Egiazarian K.: BM3D frames and variational image deblurring. IEEE Transactions on image processing 21(4), 1715-1728 (2011) . doi:10.1109/TIP.2011.2176954
5. Chen J., Li H., Chen T., Hu B., Liu S.: A Denoising Method of Remote Sensing Images Based on Improved BM3D. In: 4th International Conference on Computer Science and Application Engineering on Proceedings, pp.1-6. ACM, Zhuhai, China (2020). doi: 10.1145/3424978.3425125
6. Huang Z., Zhu Z., Wang Z., Shi Y., Fang H., Zhang Y.: DGDNet: Deep gradient descent network for remotely sensed image denoising. IEEE Geoscience and Remote Sensing Letters 20, 1-5 (2023). doi:10.1109/LGRS.2023.3241642
7. Huang Z., Wang Z., Zhu Z., Zhang Y., Fang H., Shi Y., et al.: DLRP: Learning deep low-rank prior for remotely sensed image denoising. IEEE Geoscience and Remote Sensing Letters 19, 1-5 (2022). doi: 10.1109/LGRS.2022.3167401
8. Yao Y.: Three-way decisions with probabilistic rough sets. Information sciences 180(3), 341–353 (2010). doi: 10.1016/j.ins.2009.09.021
9. Yuan K., Xu W., Li W., Ding W.: An incremental learning mechanism for object classification based on progressive fuzzy three-way concept. Information Sciences 584, 127-147 (2022). doi: 10.1016/j.ins.2021.10.058
10. Luo J., Hu M.: A bipolar three-way decision model and its application in analyzing incomplete data. International Journal of Approximate Reasoning 152, 94-123 (2023). doi: 10.1016/j.ijar.2022.10.011
11. Ying L., Miao D., Zhang Z.: A robust one-stage detector for SAR ship detection with sequential three-way decisions and multi-granularity. Information Sciences 667, 120436(2024). doi: 10.1016/j.ins.2024.120436
12. Xia S., Liu Y., Ding X., Luo Y.: Granular ball computing classifiers for efficient, scalable and robust learning. Information Sciences 483, 136-152 (2019). doi: 10.1016/j.ins.2019.01.010
13. Gao C., Tan X., Zhou J., Ding W., W Pedrycz.: Fuzzy Granule Density-Based Outlier Detection with Multi-Scale Granular Balls. IEEE Transactions on Knowledge and Data Engineering 37(5), 1182 -1197 (2025). doi: 10.1109/TKDE.2024.3525003
14. Zhang X., Yuan Z., Miao D.: Outlier detection using three-way neigh-borhood characteris-

tic regions and corresponding fusion measurement. Knowledge and Data Engineering 36(5), 2082-2095 (2024). doi: 10.1109/TKDE.2023.3312108
15. Dubois D., Prade H.: Rough fuzzy sets and fuzzy rough sets. International Journal of General System 17(2-3), 191-209 (1990). doi: 10.1080/03081079008935107
16. Tu J., Su S., Xu J.: A novel grey relational clustering model under sequential three-way decision framework. Information Sciences 663, 120248 (2024). doi: 10.1016/j.ins.2024.120248
17. Xia S., Peng D., Meng D., Zhang C., Chen Z.: Ball k-Means: Fast adaptive clustering with no bounds. IEEE transactions on pattern analysis and machine intelligence 44(1), 87–99 (2022). doi: 10.1109/TPAMI.2020.3008694
18. Haralick R., Shanmugam K., Dinstein H.: Textural features for image classification. IEEE Transactions on systems, man, and cybernetics 6, 610–621 (1973). doi: 10.1109/TSMC.1973.4309314
19. Wang Z., Liu J., Li G., Han H.: Blind2unblind: Self-supervised image denoising with visible blind spots. In: 40th the conference on computer vision and pattern recognition on Proceedings, pp. 2027-2036. IEEE, New Orleans, USA (2022). doi: 10.1109/CVPR52688.2022.00207
20. Cheng G., Zhou P., Han J.: Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. IEEE transactions on geoscience and remote sensing 54(12), 7405-7415 (2016). doi: 10.1109/TGRS.2016.2601622
21. Long, Y., Gong, Y., Xiao, Z., Liu, Q.: Accurate object localization in remote sensing images based on convolutional neural networks. IEEE transactions on geoscience and remote sensing 55(5), 2486-2498 (2017). doi: 10.1109/TGRS.2016.2645610
22. Wang C., Yeh I., Liao H.: Yolov9: Learning what you want to learn using programmable gradient information. In: 18th European conference on computer vision, pp.1-21. Springer, Milan, Italy (2024). doi: 10.1007/978-3-031-72751-1_1
23. Zhang K., Zuo W., Chen Y., Meng D.， Lei Z.: Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. IEEE Transactions on Image Processing 26(7), 3142-3155 (2017). doi: 10.1109/TIP.2017.2662206