

## Article

# Epistemic-Uncertainty-Based Divide-and-Conquer Network for Single-Image Super-Resolution

Jiaqi Yang <sup>1</sup>, Shiqi Chen <sup>1</sup>, Qi Li <sup>1</sup>, Tingting Jiang <sup>2</sup>, Yueling Chen <sup>1,\*</sup> and Jing Wang <sup>3</sup>

<sup>1</sup> State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027, China

<sup>2</sup> Research Center for Intelligent Sensing Systems, Zhejiang Laboratory, Hangzhou 311100, China

<sup>3</sup> Science and Technology on Optical Radiation Laboratory, Beijing 100854, China

\* Correspondence: chenyt@zju.edu.cn; Tel.: +86-571-87951182

**Abstract:** The introduction of convolutional neural networks (CNNs) into single-image super-resolution (SISR) has resulted in remarkable performance in the last decade. There is a contradiction in SISR between indiscriminate processing and the different processing difficulties in different regions, leading to the need for locally differentiated processing of SR networks. In this paper, we propose an epistemic-uncertainty-based divide-and-conquer network (EU-DC) in order to address this problem. Firstly, we build an image-gradient-based divide-and-conquer network (IG-DC) that utilizes gradient-based division to separate degraded images into easy and hard processing regions. Secondly, we model the IG-DC's epistemic uncertainty map (EUM) by using Monte Carlo dropout and, thus, measure the output confidence of the IG-DC. The lower the output confidence is, the more difficult the IG-DC is to process. The EUM-based division is generated by quantizing the EUM into two levels. Finally, the IG-DC is transformed into an EU-DC by substituting the gradient-based division with EUM-based division. Our extensive experiments demonstrate that the proposed EU-DC achieves better reconstruction performance than that of multiple state-of-the-art SISR methods in terms of both quantitative and visual quality.



**Citation:** Yang, J.; Chen, S.; Li, Q.; Jiang T.; Chen, Y.; Wang, J. Epistemic-Uncertainty-Based Divide-and-Conquer Network for Single-Image Super-Resolution. *Electronics* **2022**, *11*, 3809. <https://doi.org/10.3390/electronics11223809>

Academic Editor: George A. Papakostas

Received: 25 October 2022

Accepted: 17 November 2022

Published: 19 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** single-image super-resolution; epistemic uncertainty; neural networks

## 1. Introduction

Single-image super-resolution (SISR) is an important research topic in computer vision. It aims to reconstruct high-resolution (HR) images from low-resolution (LR) images. SISR has been extensively used in many fields, including information security, monitoring, medical imaging, and satellite images. However, SISR is an ill-posed problem, since multiple HR images may degenerate into one specific LR image. Numerous deep-learning SR methods have been widely developed over the last few years to establish mappings between LR and HR images. They are mainly PSNR-oriented and GAN-driven methods. PSNR-oriented methods [1–6] are trained with the MSE or L1 as loss functions and achieve excellent PSNR. Nevertheless, these losses tend to drive the super-resolution (SR) result to an average or a median of several possible SR predictions [7], causing excessive smoothing of the images. Hence, GAN-driven methods [8–13] have been proposed to address the issue of missing details. However, GAN-driven methods tend to generate pseudo-textures in the reconstructed HR image. Furthermore, by maintaining constant mapping complexity, both PSNR-oriented and GAN-driven methods fail to infer realistic details of complex structures and natural textures, ignoring the contradiction between the indiscriminate processing and the disparate difficulties in processing regions [13]. To divide and conquer the images, Wei et al. [13] proposed a network with different processing capabilities for the various components of a degraded image. Next, they divided and conquered the degraded images according to their opinions. Moreover, Wang et al. [9] proposed SFTGAN, which distinguished the processing difficulty by using HR images' category information. These division methods mentioned above are all entirely based on image information.

However, the processing difficulty is not a property of an image, but the quantification of the network processing power for different areas of images. Taking image restoration as an example, when an input image seriously degenerates, it may become challenging or impossible to deal with its restoration [14]. Further, the processing power of a model and its adaptability for coping with specific issues are critical as well [15]. Overall, quantizing the processing difficulty must encompass not only the image information, but also the model properties. Therefore, it is urgent to propose a divide-and-conquer network that follows the substance of the network processing difficulty.

Lately, the successful adoptions of Bayesian uncertainty in classification [16], segmentation [17], and camera re-localization problems [18] have shown the power of uncertainty in vision tasks. Recent progress was made by obtaining Bayesian uncertainty through dropout [16] or batch normalization [19]. The main types of uncertainty are epistemic uncertainty and aleatoric uncertainty. Epistemic uncertainty measures the output confidence of a model in processing input images. On the one hand, image information and network properties are key factors affecting output confidence. On the other hand, the lower the confidence level, the harder it is to process. Consequently, epistemic uncertainty is an appropriate measurement of the processing difficulty, as it describes how much a model is uncertain about its predictions related to the model and data. In this way, divide-and-conquer networks can localize difficult and easy processing regions more precisely by quantizing the epistemic uncertainty.

In this paper, we propose an epistemic-uncertainty-based divide-and-conquer network (EU-DC). Firstly, an image-gradient-based divide-and-conquer network (IG-DC) is built, which utilizes gradient-based division to divide. Specifically, the gradient-based division employs *Harris* to separate images into easy and hard processing areas. Secondly, we measure the output confidence of the IG-DC by applying Monte Carlo dropout in order to model the IG-DC's epistemic uncertainty map (EUM). The EUM-based division is proposed by quantizing the EUM into easy and hard processing regions. Finally, the IG-DC is transformed into an EU-DC by replacing the gradient-based division with EUM-based division.

In sum, the innovation of this paper is the proposal of a novel method for the division of processing difficulty by quantizing the epistemic uncertainty. Based on our novel division method, we further propose a divide-and-conquer SR method, which is an effective solution to the practical problem of discriminating among the difficulties in processing distinct regions. Most previous SR works ignored this problem and processed input images indiscriminately. Furthermore, some researchers only used image information (categories or gradients) to quantify the processing difficulty, but these division methods do not meet the definition of network processing difficulty. Our division considers the properties of the input image and the network's processing power for different regions of the input image. Our comprehensive experiments prove that epistemic-uncertainty-based division is reasonable and effective in quantizing the processing difficulty. Moreover, our EU-DC method is superior to the advanced SR approaches mentioned above when considering a combination of quantitative analysis and visual quality.

Our contributions can be summarized as follows:

- We introduce epistemic uncertainty in order to quantify the output confidence of the network. By utilizing the output confidence, we can clearly understand the distribution of the network's processing capabilities on the input images.
- We propose a novel division based on epistemic uncertainty, which is consistent with the substance of the processing difficulty. This division based on epistemic uncertainty accurately and reasonably distinguishes areas with different processing difficulties.
- We construct an EU-DC that divides LR images through EUM-based division and can infer clear structures and realistic textures. Extensive simulation results demonstrate that our proposed method is superior to multiple comparable state-of-the-art methods.

## 2. Related Work

In this paper, we propose an epistemic-uncertainty-based divide-and-conquer network. To improve the visual quality of the final result, an adversarial generative network is applied in our approach. Therefore, we first introduce the work related to SISR (from PSNR-oriented to GAN-driven methods). We present some of the previous divide-and-conquer approaches and their limitations for the contradiction in SISR between indiscriminate processing and disparate difficulties in processing different regions. Finally, we introduce the development of Bayesian uncertainty because we are motivated by epistemic uncertainty to propose a more reasonable and practical divide-and-conquer SISR network.

### 2.1. Single-Image Super-Resolution

Here, we review SISR methods, which can be classified into two categories: PSNR-oriented and GAN-driven methods. We also investigate specific divide-and-conquer approaches.

#### 2.1.1. PSNR-Oriented Methods

Most previous SISR networks targeted high-PSNR metrics. Dong et al. [1] initially proposed SRCNN, which introduced the convolutional neural network (CNN) to SISR and achieved superior performance to that of previous works. Kim et al. [2] designed deeper VDSR with 20 layers based on residual learning. Lim et al. [3] proposed the EDSR network by stacking modified residual blocks in which the batch-normalization layers were removed. Zhang et al. [4] proposed RCAN, which modeled the inter-dependencies between feature channels by using the channel attention mechanism and dynamically re-adjusted the weights of each channel feature. Zhang et al. [5] introduced dense connections in RDN to utilize every hierarchical feature from all of the convolutional layers. Dai et al. [6] proposed a second-order channel attention mechanism to adaptively rescale features by considering statistics higher than the first order and constructed a non-locally enhanced residual group structure to build a deep network. PSNR-oriented methods tend to drive the SR result to an average or a median of several possible SR predictions [7], causing excessive smoothing of the images.

#### 2.1.2. GAN-Driven Methods

The PSNR-oriented methods mentioned above focused on achieving high PSNR and, thus, employed the L1 or MSE loss as loss functions to measure the distance of the output results and HR images. However, the images restored by PSNR-oriented methods are always blurry. Johnson et al. [20] proposed perceptual loss to improve the visual quality of reconstructed images. Ledig et al. [8] designed SRGAN, which introduced adversarial generation networks to the field of super-resolution for the first time, and they built the first framework for generating photo-realistic HR images. Furthermore, Wang et al. [10] constructed an efficient GAN-driven framework named ESRGAN by adopting both a residual-in-residual dense block (RRDB) and perceptual loss. Soh et al. [11] designed natural manifold discrimination to make the resulting images more realistic. Ma et al. [12] proposed SPSR, which utilized image gradient information to restore clear structures. To improve the visual quality of the final result, we used the same GAN loss as that in ESRGAN [10] to optimize our proposed ED-DC. However, the GAN-driven SR methods mentioned above performed indiscriminate processing in every region with disparate difficulties, leading to a failure to infer realistic details of complex structures and textures [13].

#### 2.1.3. Divide-and-Conquer Framework

By maintaining a constant mapping complexity, neither PSNR-oriented nor GAN-driven methods can succeed in inferring the realistic details of complex structures and textures because they carry out indiscriminate processing on different regions of the input images. However, networks have distinct processing capabilities for separate areas of input

images. Some previous work was shown to be more effective for differential processing of input images [21–23]. Differential processing was also introduced into the field of SISR. Wang et al. [9] proposed SFTGAN, which divided input images based on semantic segmentation networks and employed an attention mechanism to conquer the resolution reconstruction problem. Wei et al. [13] proposed that various components' processing difficulties are diverse. Therefore, they established a division method based on the image components. It is worth noting that the processing difficulty is an abstract description of a network's processing capabilities for degraded images. Causing a differentiation in the processing power lies not only in the information of the degraded images, but also in the properties of the network. However, existing division frameworks are utterly dependent on image information, ignoring the definition of network processing difficulty.

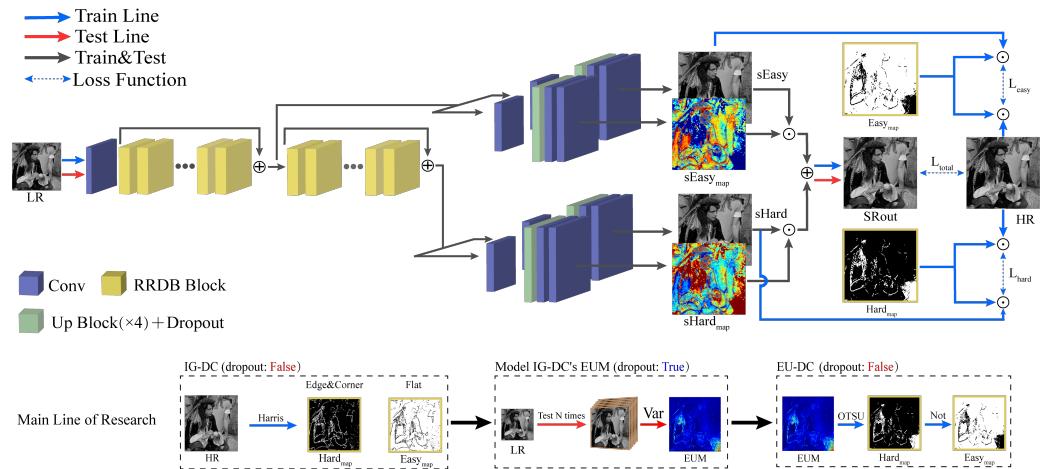
## 2.2. Bayesian Uncertainty

Understanding a model's limitations is crucial for many machine learning systems. Powerful abstract expressions from high-dimensional images for mapping to the outputs are learned by deep learning models through data-driven methods. The output results of unknown test data are often considered blindly and are believed to be reliable, which is not always true. Bayesian uncertainty plays a vital role in quantizing the output confidence of a model during testing. Output confidence helps in the decision-making process. Therefore, uncertainty is a powerful tool for any prediction and reconstruction system. Bayesian uncertainty has been proven to have advanced capacities in classification [16], segmentation [17], and camera re-localization problems [18]. In Bayesian modeling, there are two main types of uncertainty [16]. Noise inherent in the observations can be captured by aleatoric uncertainty. For instance, motion or sensor noise causes uncertainty even if a model is fed more data. On the other hand, epistemic uncertainty accounts for the uncertainty in the model parameters. Epistemic uncertainty can be explained away when given large quantities of data in machine vision and is often referred to as model uncertainty. Recent contributions utilized dropout [16] or batch normalization [19] to obtain Bayesian uncertainty. We applied the epistemic uncertainty modeling approach [16] in order to construct the divide-and-conquer SR method in this paper.

## 3. Method

### 3.1. Outline

As shown in Figure 1, the main line of the proposed method involves three steps. In the first step, inspired by [13], we construct a divide-and-conquer network structure based on image gradient division. The IG-DC divides an image into hard regions *Hard<sub>map</sub>* and easy regions *Easy<sub>map</sub>* by using *Harris* (a method that detects the gradient changes in an image). Multi-path supervision and reasonable allocation of computational overhead strategies are employed to optimize the IG-DC. In the second step, Monte Carlo dropout is introduced in order to model the IG-DC's EUM by measuring the output confidence of the IG-DC. The lower the output confidence, the more difficult the IG-DC is to process. Further, we present a division method based on epistemic uncertainty by quantizing the EUM. The division based on epistemic uncertainty is consistent with the substance of the processing difficulty because of it considers not only degraded images' features, but also the properties of the network. In the third step, we set the IG-DC as the base model. We propose the EU-DC by substituting for image-gradient-based division with EUM-based division. It is worth noting that the different steps of the main line employ the same network structure and training strategy.



**Figure 1.** The overall structure of our proposed SR method. Conv denotes a  $3 \times 3$  convolution layer. The RRDB block and up block are the same as those proposed in [10]. Up Block ( $\times 4$ ) + Dropout denotes the sequential connection of the up block ( $\times 4$ ) and dropout layer.  $\odot$  and  $\oplus$  denote element-wise multiplication and element-wise addition operations. The main line of research includes three steps. Step 1: Building the IG-DC. Step 2: Modeling the IG-DC’s output confidence to obtain division based on epistemic uncertainty. Step 3: Constructing an EU-DC by applying division based on epistemic uncertainty to the IG-DC.

### 3.2. IG-DC

Given the differences in the difficulties of reconstructing different areas, we use a divide-and-conquer network structure to build the IG-DC. Motivated by [13], the IG-DC measures the processing difficulty by using *Harris* to detect gradient changes in images. Specifically, as shown in Figure 1, an HR image is divided into three components by *Harris*. The edge and corner components are identified as the hard regions. Meanwhile, the flat component is defined as a simple region. Moreover, we adopt multi-path supervision and strategies for reasonable allocation of computational overhead to optimize the IG-DC.

#### 3.2.1. Multi-Path Supervision

To facilitate the feature learning in the IG-DC from easy to hard, the IG-DC built easy and hard branches to overcome regions with various processing difficulties. We define the results of the easy routing as  $sEasy$  and  $sEasy_{map}$ . Meanwhile,  $sHard$  and  $sHard_{map}$  are the outputs of the hard branch. The overall result  $SRout$  is obtained by using weighted fusion of the results of the two routings:

$$SRout = sEasy \odot sEasy_{map} + sHard \odot sHard_{map}, \quad (1)$$

Different loss functions are employed to supervise the multi-path SR results.  $SRout$  is supervised by utilizing a combination loss  $L_G$  comprising the pixel loss  $L_{pix}$ , perceptual loss  $L_{per}$  [20], and GAN loss  $L_{gan}$ . The definitions of  $L_{pix}$ ,  $L_{per}$ , and  $L_{gan}$  are:

$$L_{pix} = ||SR - HR||_1, \quad (2)$$

$$L_{per} = ||\Phi_i(SR) - \Phi_i(HR)||_1, \quad (3)$$

where  $\Phi_i(\cdot)$  denotes the  $i_{th}$ -layer output of the VGG [24] model.

$$L_{gan} = -E_{SR}[\log(D(SR, HR))] - E_{HR}[\log(1 - D(HR, SR))], \quad (4)$$

Thus, the combination loss  $L_G$  is shown below:

$$L_G = \alpha_1 L_{pix} + \beta_1 L_{per} + \gamma_1 L_{gan}, \quad (5)$$

where  $\alpha_1$ ,  $\beta_1$ , and  $\gamma_1$  denote the trade-off hyperparameters of different losses. The loss for the easy branch is formulated as:

$$L_{easy} = ||Easy_{map} \odot sEasy - Easy_{map} \odot HR||_1, \quad (6)$$

$$L_{hard} = ||Hard_{map} \odot sHard - Hard_{map} \odot HR||_1, \quad (7)$$

Finally, the loss function for the IG-DC's generator is expressed as:

$$L_{total} = \alpha_2 L_{easy} + \beta_2 L_{hard} + \gamma_2 L_G, \quad (8)$$

where  $\alpha_2$ ,  $\beta_2$ , and  $\gamma_2$  denote the trade-off hyperparameters of different branches.

The IG-DC's discriminator network is a VGG128 model [24], and the discriminator loss is formulated as:

$$L_D = -E_{SR}[\log(D(HR, SR))] - E_{HR}[\log(1 - D(SR, HR))], \quad (9)$$

The IG-DC's discriminator and generator are optimized through adversarial learning. Through multi-path supervision, we guide the network in emphatically learning processing-difficulty-attentive masks, with  $Hard_{map}$  and  $Easy_{map}$  providing guidance from the gradient information in the HR images. In other words,  $sEasy$  focuses more on the generation of simple regions because of the supervision from Equation (6), which is similar to that of  $sHard$ . Furthermore, to produce  $SRout$  with a higher quality, it is proper for the direction of optimization for  $sHard_{map}$  and  $sEasy_{map}$  to be regional distributions of  $Hard_{map}$  and  $Easy_{map}$ , respectively. The IG-DC's generator produces processing-difficulty-attentive masks and intermediate SR predictions. Therefore, relatively independent branches supervise the various regions, making it possible to perform differential processing.

### 3.2.2. Reasonable Allocation of Computational Overhead

To reasonably allocate the computational overhead of the network, we build the generator of the IG-DC with a stacked architecture comprising 20 RRDBs, making it beneficial for adjusting our emphases on different regions. The main factors limiting the overall quality of reconstruction are the areas that are challenging to reconstruct. Therefore, putting more attention into extracting features in problematic areas is, theoretically, a more reasonable option. We exploit the first part of the stacked architecture to reconstruct accessible areas and the rest to inherit the first part's output for restoring complex areas. At this point, we carry out the reasonable allocation of computational overhead in the feature extraction stage. We utilize the feature extraction stage and the backpropagation phase to achieve discriminate processing. Distinct values are configured for  $\alpha_2$  and  $\beta_2$  in Equation (8) to modify the proportions of the original gradients of the different branches, which is a method for implicitly adjusting the allocation of computational overhead.

### 3.3. Modeling the IG-DC's EUM

The processing difficulty distribution map of the IG-DC is based on gradient changes in HR images, which is not consistent with the substance of network processing difficulty.

A reasonable processing difficulty division method should have two characteristics. On the one hand, the input image information and model properties play a decisive role. On the other hand, it should have a stable relationship of transformation with the processing difficulty. The model's output confidence is a good choice because it falls perfectly in line with these two characteristics. To measure the model's output confidence, we introduce epistemic uncertainty into SISR. Compared to modeling the EUM by using batch normalization, adopting dropout as a Bayesian approximation alleviates the computational consumption and accuracy degradation that may be caused by the uncertainty characterization process in deep learning models. Thus, we model the EUM of the IG-DC by using the Monte Carlo dropout. Specifically, the status of the dropout is configured to "true". We

utilize the dropout to randomly sample the image features after the up block ( $\times 4$ ) and then reconstruct the HR image by using the sampled features.

In training, the network structure and training strategy remain the same as those in the IG-DC. In the test procedure, we enter an LR image as an input  $N$  times and obtain a set of different outputs  $[SRout_1, SRout_2, \dots, SRout_{N-1}, SRout_N]$ . The EUM is formulated as the variance of the outputs:

$$EUM = \text{norm}\left(\frac{\sum_i SRout_i^2}{N} - \left(\frac{\sum_i SRout_i}{N}\right)^2\right), \quad (10)$$

where  $\text{norm}(x) = x/x_{max}$  is the operation of normalization.

Finally, we obtain the IG-DC's EUM for each image in the training dataset.

### 3.4. EU-DC

To build a divide-and-conquer network that is consistent with the network processing difficulty, we transform the IG-DC into the EU-DC by substituting the image-gradient-based division with EUM-based division. On the one hand, the EUM considers the properties of both the input image and the network. On the other hand, the value of the EUM has 256 levels; thus, the difficulty of image processing is divided more accurately and finely. Further, the EUM-based division is constructed with the EUM as the prior information of the IG-DC's processing difficulty distribution. Specifically, we quantify the EUM according to two levels to define  $Easy_{map}$  and  $Hard_{map}$ . As shown in Figure 1, the EUM is binarized by OTSU [25] to get the hard areas, i.e.,  $Hard_{map} = \text{OTSU}(EUM)$ . The  $Easy_{map}$  can be obtained by reversing  $Hard_{map}$ , i.e.,  $Easy_{map} = 1 - Hard_{map}$ .

Before the training of the EU-DC, we utilized the Monte Carlo dropout to model the epistemic uncertainty on all of Div2K's data. To obtain the values of  $Hard_{map}$  and  $Easy_{map}$  required for network training, we employed OTSU to quantify the epistemic uncertainty map into two levels. Finally, we applied  $Hard_{map}$  and  $Easy_{map}$  to guide the network in restoring the complicated regions and accessible regions, respectively.

In the training procedure, by obtaining the IG-DC's EUM for each image in the second step of our main line, we no longer need the dropout to model the EUM; thus, the dropout is turned off. The EU-DC utilizes the same dataset, loss function, and training strategy as those in the ID-DC to perform supervised learning. Compared with the IG-DC, the EUM is an additional input for the SR model because we need to use the EUM-based division to transform the EUM into  $Hard_{map}$  and  $Easy_{map}$ . Through the EUM's guidance, the EU-DC accurately allocates more computing power to areas in which IG-DC does not process well and improves the overall processing performance in these regions. In other words,  $Hard_{map}$  according to the EUM is where the IG-DC does not perform well. The EU-DC pays attention to  $Hard_{map}$  to obtain more excellent overall capabilities. During the test, we no longer need to enter the EUM and only input the degraded image because the EUM's only role is to guide the supervised learning of EU-DC, and generating  $SRout$  does not require the EUM. The EU-DC can reconstruct HR images and predict the processing-difficulty-attentive masks, accessible processing areas  $sEasy_{map}$ , and complex processing areas  $sHard_{map}$ .

In summary, we transform the IG-DC into the EU-DC by replacing the gradient-based division method with the EUM-based division method. EU-DC is an upgraded version that improves upon the IG-DC's imperfections.

## 4. Experiment

To evaluate our technique, we carried out a comprehensive set of experiments with the aim of answering the following two questions:

- What are the superiorities of the proposed EU-DC SR model? The answer to this question is based on its characteristics, including the allocation of computational overhead, the analysis of the EUM, and the division method.
- Is the proposed reconstruction solution superior to the state-of-the-art SR methods when comparison the combination of a quantitative analysis and the visual quality?

The answer to this question is based on a comprehensive comparison with other SR methods. We analyzed objective metrics, visual quality, model parameters, and running times.

#### 4.1. Implementation Details

The IG-DC and EU-DC were trained with a training set (800 images) from DIV2K [26]. Both of the hyperparameters—the block size and window size for Harris—in the IG-DC were set to 3. The free parameter in the corner detection equation for Harris was to 0.04. Moreover, the training set was processed with a scaling factor of  $\times 4$  between the LR and HR images. The input LR images were obtained by down-sampling their GT images by using the bicubic method [27]. We set the batch size to 16. The spatial size of the cropped HR patch was  $128 \times 128$ . The Adam optimizer [28] was used, in which  $\beta_1$  was 0.9 and  $\beta_2$  was 0.99. We set the learning rates to  $1 \times 10^{-4}$  for both the generator and the discriminator and reduced them to half at 50 k, 100 k, 200 k, and 300 k iterations. The hyperparameters in Equation (5), namely,  $\alpha_1$ ,  $\beta_1$ , and  $\gamma_1$ , were set to 0.1, 1, and 0.005, respectively.  $\alpha_2$ ,  $\beta_2$ , and  $\gamma_2$  in Equation (8) were set to 1, 4, and 1, respectively. We set the sampling rate of the dropout to 0.2. Our stacked architecture in the IG-DC included 20 RRDB blocks. We allocated 5 and 15 RRDB blocks for the easy and hard regions, respectively. Before training the EU-DC, we first obtained the IG-DC's EUM for each image in the training set (800 images) from DIV2K. For the test of the EU-DC, we used four benchmark datasets, namely, Set5 [29], Set14 [30], BSDS100 [31], and Urban100 [32].

#### 4.2. Superiority Analysis

To demonstrate the superiority of the EU-DC, we performed comprehensive studies on our method. As shown in Table 1, the base model was the IG-DC.  $N_a$  and  $N_b$  denote the  $N_a$  basic RRDB blocks that reconstructed simple areas and the other  $N_b$  basic RRDB blocks that served to restore difficult areas.  $\alpha_2$  and  $\beta_2$  in Table 2 are identical to those in Equation (8). All results from the reconstructed degraded images in the Set14 dataset were comprehensively evaluated with objective and visual metrics. The objective metrics included the PSNR and SSIM [33], and the visual metrics included the NIQE [34] and LPIPS [35]. The higher the indicators of the PSNR and SSIM, the better the quantitative quality of the reconstructed image. The lower the values of the NIQE and LPIPS, the better the visual quality of the reconstructed image.

**Table 1.** A study of the RRDB allocation schemes in the Set14 dataset.

Scheme	PSNR	SSIM	NIQE	LPIPS
IG-DC + 5 & 15	26.463	0.7800	3.827	0.1419
IG-DC + 10 & 10	26.453	0.7751	3.811	0.1417
IG-DC + 15 & 5	26.458	0.7760	3.843	0.1429
EU-DC + 5 & 15	27.538	0.8161	3.621	0.1309
EU-DC + 10 & 10	27.144	0.8058	3.501	0.1257
EU-DC + 15 & 5	27.504	0.8094	3.602	0.1324

**Table 2.** A study of the loss weights in the Set14 dataset.

Scheme $\alpha_2$ & $\beta_2$	PSNR	SSIM	NIQE	LPIPS
1 & 1	27.538	0.8161	3.621	0.1309
1 & 2	27.537	0.8159	3.617	0.1306
1 & 4	27.535	0.8158	3.612	0.1297
1 & 6	27.531	0.8142	3.619	0.1307
1 & 8	27.109	0.8029	3.629	0.1337

#### 4.2.1. Allocation of Computational Overhead

In order to prove that our proposed network structure can reasonably allocate computing power, we conducted experiments on control variables for module allocation. Both the IG-DC and EU-DC divided an input image into two difficulty levels, resulting in the final reconstruction being affected by the restoration qualities of regions with distinct ratings. Discriminate processing was carried out in the feature extraction and backpropagation phases. We demonstrated the roles of these two stages by using control variates.  $\alpha_2 = 1$  and  $\beta_2 = 1$  were our initial conditions. It is worth emphasizing that the EU-DC and its corresponding IG-DC adopted the same network structure and loss functions, regardless of the allocation scheme employed.

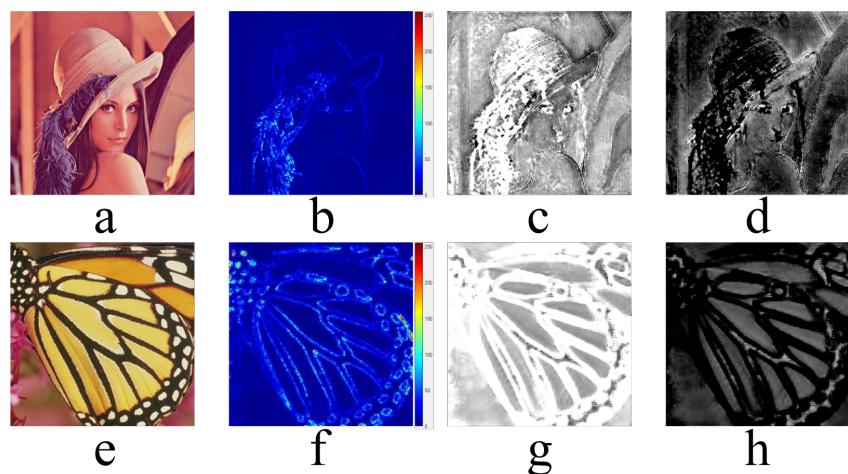
To demonstrate the essential effect of the stacked structures in obtaining excellent overall performance, we experimented with RRDB allocation schemes under the conditions of  $\alpha_2 = 1$  and  $\beta_2 = 1$ . As shown in Table 1, we adopted three different schemes for allocation of computational overhead. The distribution of “5 & 15” means that more computational overhead was concentrated on challenging areas. By contrast, “15 & 5” denotes that we paid more attention to reconstructing simple regions. “10 & 10” is a balanced option. Compared to “15 & 5”, “5 & 15” surpasses it in all indicators, regardless of if the IG-DC or EU-DC is used. Between “5 & 15” and “10 & 10”, “5 & 15” obtains a better quantitative quality, whereas “10 & 10” is superior in visual quality. It is worth noting that “5 & 15” surpasses “10 & 10” very much in objective quality, but “10 & 10” does not surpass “5 & 15” very much in visual quality. Therefore, “5 & 15” is an excellent option for the IG-DC and EU-DC.

Moreover, to prove the effect of modifying the loss weights on the allocation of the computational power, an experiment on the loss weights was performed with the best RRDB allocation scheme, “5 & 15”. We fixed the value of  $\alpha_2$  to 1 and modified the various branches’ loss weights by adopting values from the list [1, 2, 4, 6, 8] for  $\beta_2$ . As shown in Table 2, we transformed  $\beta_2$  from 1 to 4. The objective indicators slightly decreased, while the visual indicators significantly improved, especially in the NIQE. Further, by continuing to raise the value of  $\alpha_2$ , our emphasis was heavily unbalanced, causing all indicators to decline to varying degrees. Consequently, setting  $\beta_2$  to 4 was a great choice for improving the all-around performance, and the loss weights affected the distribution of the computing power and balanced the overall effect of reconstruction.

To put it in a nutshell, the EU-DC reasonably allocated computational overhead in both the feature extraction and backpropagation phases, and it realized an exceptional balance between quantitative quality and visual performance.

#### 4.2.2. Analysis of the EUM

In order to analyze the rationality of quantizing the processing difficulty with the EUM, we visually analyzed it. As shown in Figure 2, we utilized MATLAB’s colormap to visualize the IG-DC’s EUM for a more intuitive understanding. We first analyze Figure 2a. The information in “lenna” contains fine textural details and edge structures. Figure 2b shows the EUM of “lenna”. The large numerical values are mainly concentrated in the hair texture area, which indicates that it is challenging for the IG-DC to reconstruct complex textures. This is consistent with the fact that GAN-driven SR methods tend to create pseudo-textures in complex regions. Figure 2e includes more flat regions, and Figure 2f mainly positions the edge structure, which demonstrates that the EUM dynamically adapts to input images with different characteristics. In addition, the EUM quantifies the processing difficulty into 256 levels, which is more refined than the gradient-based division method. To illustrate that the EUM can effectively guide the EU-DC in dividing and conquering input images, the processing-difficulty-attentive masks ( $sEasy_{map}$  and  $sHard_{map}$ ) are also presented in Figure 2. Figure 2c accurately locates the area with a high value in the EUM, and it is regionally complementary to Figure 2d. Figure 2g,h also complement each other in certain regions, which proves that the network can effectively carry out regional supervised learning.



**Figure 2.** The EUM of the IG-DC. (a,e) “lenna” from Set14 and “butterfly” from Set5, (b,f) lenna’s EUM and butterfly’s EUM from the ID-DC, (c,g) lenna’s  $sEasy_{map}$  and butterfly’s  $sEasy_{map}$  from the EU-DC, (d,h) lenna’s  $sHard_{map}$  and butterfly’s  $sHard_{map}$  from the EU-DC.

To sum up, on the one hand, the regions with higher values in the EUM were, indeed, strongly correlated with areas that were difficult to handle. On the other hand, the EUM was able achieve more refined division. Consequently, the EUM is a reasonable and precise program for quantifying processing difficulty.

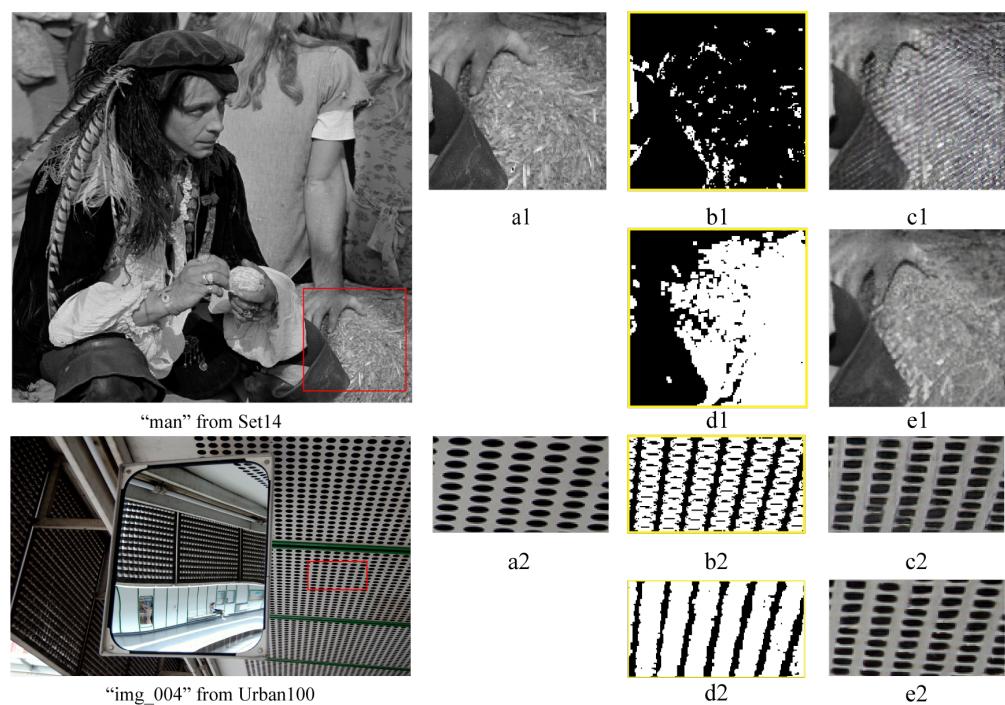
#### 4.2.3. Division Method

We performed a comprehensive experimental analysis to demonstrate that epistemic-uncertainty-based division is more reasonable and effective in quantizing the processing difficulty than image-gradient-based division. As shown in Figure 3, on the one hand,  $Hard_{map}$  varied widely for the IG-DC and EU-DC. The low numerical distribution in Figure 3(b1) indicates that the IG-DC model considered the ROI of “man” to be easily reconstructed for the network. By contrast, the HR image in Figure 3(c1) that was restored by the ID-DC was full of pseudo-textures, which indicated that gradient-based difficulty division methods do not accurately locate the distributions of actual processing difficulties. Meanwhile, Figure 3(d1) shows that the EU-DC accurately located difficult processing areas. The HR results in Figure 3(e1) show that a natural texture was reconstructed. On the other hand,  $Hard_{map}$  was similar for the IG-DC and EU-DC. As shown in Figure 3(b2,d2), Figure 3(b2) is more regular and has many fine structures, which is unreasonable because images with degraded resolutions have often lost these tiny features. Consequently, distortions appear in the reconstructed HR image in Figure 3(c2). Figure 3(d2) is relatively flat and can completely cover delicate structures. The results restored by the EU-DC in Figure 3(e2) are well structured.

Based on the above analysis, EUM-based division is more reasonable and adequate than image-gradient-based division because EUM-based division considers not only the input image information, but also the network’s properties for a particular degraded image.

In addition, as shown in Table 1, no matter what allocation of computational overhead scheme is adopted, under conditions in which the network structure and training strategy are identical, the EU-DC surpasses the IG-DC by a large margin for all indicators. The improvement in the indicators further illustrates the superiority of the EUM-based division method.

In conclusion, the EU-DC is an advanced SISR model because it can accurately and effectively divide the difficulty levels and reasonably allocate the computational overhead for different ratings of areas.



**Figure 3.** Comparison the IG-DC and EU-DC. (a1,a2) man's original and img\_004's original of ROI, (b1,b2) man's *Hard<sub>map</sub>* and img\_004's *Hard<sub>map</sub>* defined by the IG-DC, (c1,c2) man's result and img\_004's result restored by the IG-DC, (d1,d2) man's *Hard<sub>map</sub>* and img\_004's *Hard<sub>map</sub>* defined by the EU-DC, and (e1,e2) man's result and img\_004's result restored by the EU-DC.

#### 4.3. Quantitative Analysis

##### 4.3.1. Quantitative Evaluation

To demonstrate the superiority of the proposed method, we compared it with a variety of recently proposed SR networks, including ESRGAN [10], NatSR [11], SPSR [12], and ATG [36]. We used the PSNR, SSIM [33], NIQE [34], and LPIPS [35] as evaluation metrics. Overall, our EU-DC method was able to achieve a comparable or superior performance with respect to its existing counterparts.

In Table 3, we present the results of recent advanced super-resolution network methods on the Set5, Set14, BSDS100, and Urban100 datasets. In comparison with ESRGAN, our approach surpassed it by a large margin for all indicators. NatSR obtained excellent objective quality in terms of the PSNR and SSIM metrics. However, it obtained unsatisfactory results for the visual metrics, suggesting that NatSR tends to produce relatively blurry results with a high PSNR compared to the results of other perceptually driven methods. Our method surpassed NatSR in the objective metrics and maintained a superior visual quality. SPSR achieved good results for the visual indicators by introducing image gradient information in order to suppress image distortion. At the same time, our method achieved higher scores on all indicators than SPSR did in the objective evaluation. ATG obtained well-rounded metrics, especially on the Set5 dataset, but our method obtained better indicator scores when considering all of the metrics for all datasets.

Therefore, our EU-DC method comprehensively achieved excellent indicator scores on all test datasets, and it is superior to the recent SR methods with which it was compared in terms of both quantitative and visual quality indicators.

**Table 3.** Comparison with state-of-the-art SR methods on the Set5, Set14, BSDS100, and Urban100 datasets. The best performance is highlighted in red (best) and blue (second best).

Dataset	Metric	Bicubic	ESRGAN	NatSR	SPSR	ATG	EU-DC
<b>Set5</b>	NIQE	8.4927	5.1279	5.6569	<b>4.6126</b>	5.7913	<b>3.7154</b>
	LPIPS	0.3440	0.0740	0.0945	0.0711	<b>0.0667</b>	<b>0.0689</b>
	PSNR	28.385	30.278	30.971	30.382	<b>31.532</b>	<b>31.032</b>
	SSIM	0.8249	0.8666	<b>0.8807</b>	0.8635	<b>0.8876</b>	0.8773
<b>Set14</b>	NIQE	7.7208	<b>3.8098</b>	3.8807	3.8648	4.0568	<b>3.6117</b>
	LPIPS	0.4419	0.1378	0.1769	0.1327	<b>0.1312</b>	<b>0.1297</b>
	PSNR	26.084	26.337	<b>27.512</b>	26.635	27.399	<b>27.535</b>
	SSIM	0.7849	0.7811	<b>0.8146</b>	0.7939	0.8112	<b>0.8158</b>
<b>BSDS100</b>	NIQE	7.7050	3.5099	3.9584	<b>3.3902</b>	3.9852	<b>3.3149</b>
	LPIPS	0.5252	0.1782	0.2099	0.1758	<b>0.1598</b>	<b>0.1713</b>
	PSNR	25.944	25.707	<b>26.447</b>	25.503	<b>26.459</b>	26.342
	SSIM	0.6686	0.6658	0.6849	0.6592	<b>0.6932</b>	<b>0.6787</b>
<b>Urban100</b>	NIQE	7.3326	3.8796	3.9405	<b>3.8651</b>	4.1292	<b>3.8062</b>
	LPIPS	0.4742	0.1291	0.1503	0.1270	<b>0.1241</b>	<b>0.1237</b>
	PSNR	23.130	24.701	25.457	24.792	<b>25.559</b>	<b>25.674</b>
	SSIM	0.9009	0.9456	0.9505	0.9481	<b>0.9557</b>	<b>0.9571</b>

#### 4.3.2. Model Size and Running Times

We analyzed the computational complexity of the EU-DC and compared it with those of other advanced SR models. We first show the model parameter sizes of the EU-DC and the other advanced SR models in Table 4. In addition, we recorded the running time that it took for each method to test the Urban100 dataset.

**Table 4.** A study of the model size and running time. “M” denotes MByte.

	ESRGAN	NatSR	SPSR	ATG	EU-DC
Model parameters	63.94 M	53.67 M	94.93 M	63.91 M	58.51 M
Running times	28.14 s	24.98 s	28.45 s	10.70 s	13.16 s

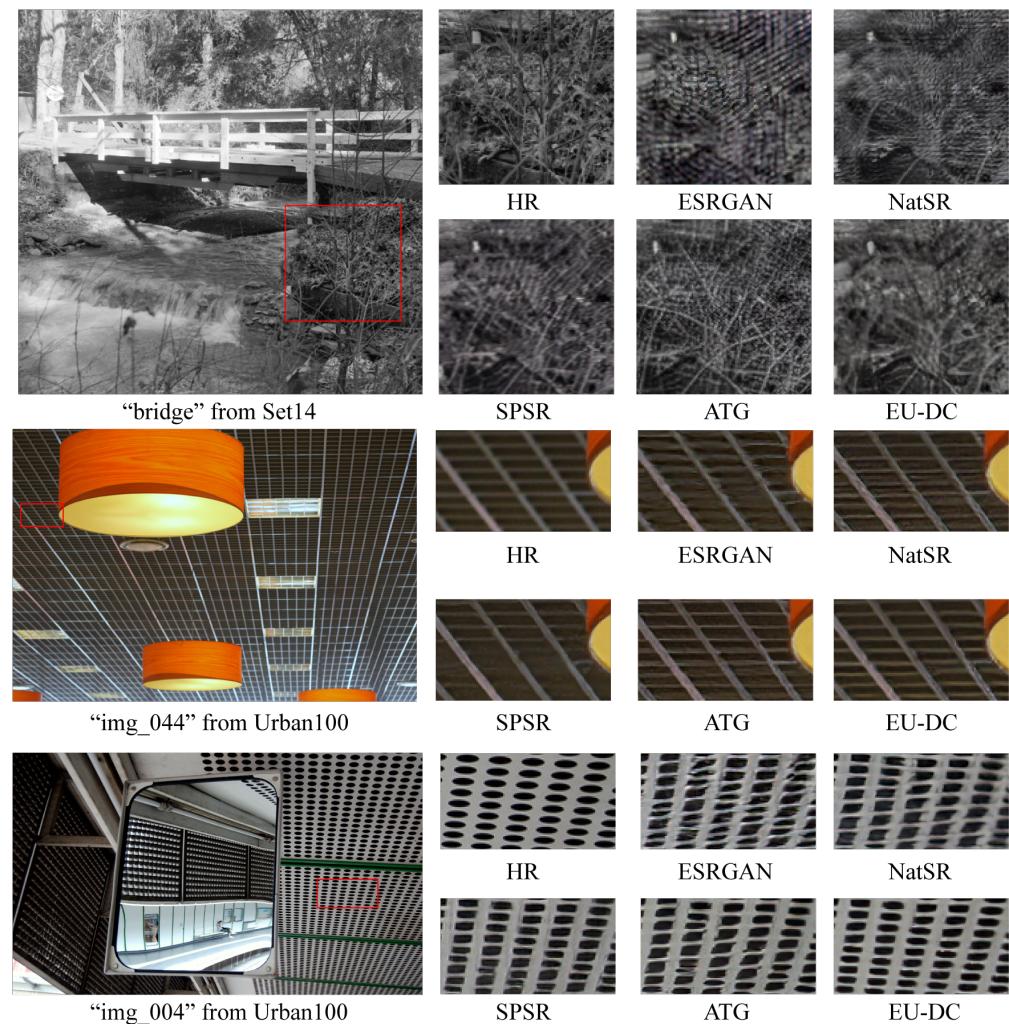
**Model Size:** The parameter size of NatSR was the most lightweight, but NatSR lacked superior performance. The EU-DC’s parameter size was secondary, and there was only a small gap between the EU-DC and NatSR. In terms of the comprehensive performance comparison, the EU-DC far exceeded NatSR. The model parameters of SPSR were the largest, but it did not obtain the best overall indicator scores. ATG achieved slightly inferior metrics than the EU-DC, but it used more model parameters than the EU-DC. It is worth noting that ESRGAN, SPSR, and the EU-DC all employed stacked RRDB structures to extract image features, but the EU-DC built the best-performing model by utilizing the minimum number of RRDBs. Therefore, the EU-DC is more efficient in utilizing computational resources to achieve better reconstruction.

**Running times:** We evaluated the proposed and start-of-the-art methods mentioned above on the Urban100 dataset. We performed all experiments on a GeForce RTX 2080Ti with 11 GB of memory. The running times are shown in Table 4. The running times of the proposed method were slightly inferior to those of ATG. In contrast, our approach was faster than ESRGAN, NatSR, and SPSR, which shows the EU-DC’s advantages in terms of computational complexity.

In conclusion, the EU-DC utilizes computational resources effectively, and its computational complexity is lower than those of other advanced SR models.

#### 4.4. Visual Quality Comparison

We conducted visual quality comparisons to demonstrate the superiority of our approach more intuitively. In addition to the quantitative analysis conducted above, Figure 4 depicts a visualization of the results of our EU-DC method and the methods with which it was compared. First, taking bridge.png in the Set14 dataset as an example, our EU-DC restored more realistic tree textures than the other SR models did, thus making the rebuilt image more natural. On the contrary, the other models produce plenty of pseudo-texture and distortions.



**Figure 4.** Visual results of different GAN-driven methods. Our proposed method reduced the distortion in the image and produced more natural textures and clearer structures than the other SR methods did.

Another case is img\_044.png from the Urban100 dataset. Restoring small periodic structures and making them easy to identify is challenging. Therefore, previously proposed SR approaches, from ESRGAN to ATG, have failed to address the problem. In their results, these tiny structures still have varying degrees of blur, so it is challenging to define the edges of each fine detail. Our EU-DC method can restore more structural minutiae, thus making finer structures distinguishable.

As for img\_004.png in the Urban100 dataset, restoring tiny and regular appearances is notably tricky. The ellipses in img\_004.png recovered by ESRGAN and NatSR were massively distorted. The results restored by SPSR and ATG improved to some extent. Compared with the mentioned methods, the EU-DC reconstructed regular ellipses with little distortion.

In summary, compared with many recently proposed advanced SR models, our proposed EU-DC model recovers discernible structures and natural textures, resulting in excellent quantitative and visual quality.

## 5. Conclusions

In this paper, we proposed a novel EU-DC model that achieved the restoration of HR images with clear structures and realistic textures. First, we employed a divide-and-conquer framework to build an IG-DC, which not only progressively facilitated the model's feature learning, but also reasonably allocated computational overhead. Next, we modeled the EUM of the IG-DC by using dropout. Finally, the IG-DC was transformed into an EU-DC by substituting the image-gradient-based division method with EUM-based division. Extensive experiments demonstrated that epistemic-uncertainty-based division is reasonable and effective in quantizing the processing difficulty. The EU-DC greatly surpassed the IG-DC in all evaluations with objective indexes, especially the PSNR (an increase of at least 0.69 dB, as shown in Table 1). Moreover, our EU-DC method achieves excellent comprehensive indicator scores for all test datasets and alleviated geometric distortions that commonly exist in the SR results of perceptually driven methods. In conclusion, the EU-DC is comprehensively superior to other advanced SR approaches in terms of the combination of quantitative analysis and visual quality.

## 6. Future Work

Although our method achieved good performance, we did not take the best advantage of the EUM. The EUM can divide the processing difficulty into 256 ratings, but here, we quantified it into simple and complicated levels without much thought. Making more effective use of the EUM will be our future work. In addition, at present, we fixed the allocation ratios for the operation modules. In the future, we will design the ratios as a training parameter so that the network can automatically adjust the balance according to the input image. Our proposed divide-and-conquer framework is universal for image restoration. The framework is especially beneficial for non-uniform degradation of image restoration because the EUM can accurately locate areas in which the degradation is severe. Therefore, we will explore the potential of this method and expand its application to other image restoration tasks, such as dehazing or deblurring. In addition, pseudo-textures and distortions often appear in the results restored with generative denoising methods in real-world scenarios. We will try to address this problem by utilizing uncertainty caused by noise as prior information.

**Author Contributions:** Conceptualization, J.Y.; Methodology, J.Y. and S.C.; Software, J.Y. and T.J.; Investigation, J.Y.; Writing—original draft, J.Y.; Writing—review & editing, S.C.; Visualization, Q.L.; Supervision, Q.L.; Project administration, Y.C.; Funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the Equipment Pre-Research Key Laboratory Fund (612102Y030306), the Key Research Project of Zhejiang Lab (2021MH0AC01) and the National Natural Science Foundation of China (62275229).

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Meijuan Bian from the facility platform of optical engineering of Zhejiang University for instrument support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 295–307. [[CrossRef](#)] [[PubMed](#)]
2. Kim, J.; Lee, J.K.; Lee, K.M. Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, CA, USA, 26 June–1 July 2016; pp. 1646–1654.

3. Lim, B.; Son, S.; Kim, H.; Nah, S.; Mu Lee, K. Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Hawaii, HI, USA, 21–26 July 2017; pp. 136–144.
4. Zhang, Y.; Li, K.; Li, K.; Wang, L.; Zhong, B.; Fu, Y. Image super-resolution using very deep residual channel attention networks. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 286–301.
5. Zhang, Y.; Tian, Y.; Kong, Y.; Zhong, B.; Fu, Y. Residual dense network for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 2472–2481.
6. Dai, T.; Cai, J.; Zhang, Y.; Xia, S.T.; Zhang, L. Second-order attention network for single image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 11065–11074.
7. Lehtinen, J.; Munkberg, J.; Hasselgren, J.; Laine, S.; Karras, T.; Aittala, M.; Aila, T. Noise2Noise: Learning image restoration without clean data. *arXiv* **2018**, arXiv:1803.04189.
8. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, HI, USA, 21–26 July 2017; pp. 4681–4690.
9. Wang, X.; Yu, K.; Dong, C.; Loy, C.C. Recovering realistic texture in image super-resolution by deep spatial feature transform. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 606–615.
10. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Change Loy, C. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018; pp. 63–79.
11. Soh, J.W.; Park, G.Y.; Jo, J.; Cho, N.I. Natural and realistic single image super-resolution with explicit natural manifold discrimination. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 8122–8131.
12. Ma, C.; Rao, Y.; Cheng, Y.; Chen, C.; Lu, J.; Zhou, J. Structure-preserving super resolution with gradient guidance. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, DC, USA, 14–19 June 2020; pp. 7769–7778.
13. Wei, P.; Xie, Z.; Lu, H.; Zhan, Z.; Ye, Q.; Zuo, W.; Lin, L. Component divide-and-conquer for real-world image super-resolution. In Proceedings of the European Conference on Computer Vision, Online, 23–28 August 2020; pp. 101–117.
14. Chen, S.; Feng, H.; Pan, D.; Xu, Z.; Li, Q.; Chen, Y. Optical Aberrations Correction in Postprocessing Using Imaging Simulation. *ACM Trans. Graph. (TOG)* **2021**, *40*, 1–15. [[CrossRef](#)]
15. Chen, S.; Feng, H.; Gao, K.; Xu, Z.; Chen, Y. Extreme-quality computational imaging via degradation framework. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 10–17 October 2021; pp. 2632–2641.
16. Kendall, A.; Gal, Y. What uncertainties do we need in bayesian deep learning for computer vision? *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5580–5590.
17. Huang, P.Y.; Hsu, W.T.; Chiu, C.Y.; Wu, T.F.; Sun, M. Efficient uncertainty estimation for semantic segmentation in videos. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 520–535.
18. Kendall, A.; Cipolla, R. Modelling uncertainty in deep learning for camera relocalization. In Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 16–21 May 2016; pp. 4762–4769.
19. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.
20. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 10–16 October 2016; pp. 694–711.
21. Farghaly, M.; Mansour, R.F.; Sewisy, A.A. Two-stage deep learning framework for sRGB image white balance. *Signal Image Video Process.* **2022**, *1*–8. [[CrossRef](#)]
22. Mansour, R.F.; Escorcia-Gutierrez, J.; Gamarra, M.; Villanueva, J.A.; Leal, N. Intelligent video anomaly detection and classification using faster RCNN with deep reinforcement learning model. *Image Vis. Comput.* **2021**, *112*, 104229. [[CrossRef](#)]
23. Li, L.; Sun, L.; Xue, Y.; Li, S.; Huang, X.; Mansour, R.F. Fuzzy multilevel image thresholding based on improved coyote optimization algorithm. *IEEE Access* **2021**, *9*, 33595–33607. [[CrossRef](#)]
24. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
25. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
26. Agustsson, E.; Timofte, R. Ntire 2017 challenge on single image super-resolution: Dataset and study. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 126–135.
27. Keys, R. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Process.* **1981**, *29*, 1153–1160. [[CrossRef](#)]
28. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
29. Bevilacqua, M.; Roumy, A.; Guillemot, C.; Alberi-Morel, M.L. Low-Complexity Single-Image Super-Resolution Based on Nonnegative Neighbor Embedding. In Proceedings of the British Machine Vision Conference, Guildford, UK, 3–7 September 2012; pp. 135.1–135.10.
30. Zeyde, R.; Elad, M.; Protter, M. On single image scale-up using sparse-representations. In Proceedings of the International Conference on Curves and Surfaces, Avignon, France, 24–30 June 2010; pp. 711–730.

31. Martin, D.; Fowlkes, C.; Tal, D.; Malik, J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In Proceedings of the Eighth IEEE International Conference on Computer Vision, ICCV 2001, Vancouver, BC, Canada, 7–14 July 2001; Volume 2, pp. 416–423.
32. Huang, J.B.; Singh, A.; Ahuja, N. Single image super-resolution from transformed self-exemplars. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, FL, USA, 7–12 June 2015; pp. 5197–5206.
33. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
34. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [[CrossRef](#)]
35. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 586–595.
36. Jo, Y.; Oh, S.W.; Vajda, P.; Kim, S.J. Tackling the ill-posedness of super-resolution through adaptive target generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Online, 19–25 June 2021; pp. 16236–16245.