Article

A Novel Underwater Image Enhancement Using Optimal Composite Backbone Network

Yuhan Chen [1], Qingfeng Li [2], Dongxin Lu [2], Lei Kou [3], Wende Ke [1,*], Yan Bai [1] and Zhen Wang [1]

1 Department of Mechanical and Energy Engineering, Southern University of Science and Technology, Shenzhen 518055, China; 12132246@mail.sustech.edu.cn (Y.C.); 11930359@mail.sustech.edu.cn (Y.B.); 12032414@mail.sustech.edu.cn (Z.W.)
2 Health Management System Engineering Center, School of Public Health, Hangzhou Normal University, Hangzhou 311121, China; 1030779697@qq.com (Q.L.); zn19980@sina.com (D.L.)
3 Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences), Qingdao 266075, China; koulei1991@qlu.edu.cn
* Correspondence: kewd@sustech.edu.cn

Abstract: Continuous exploration of the ocean has made underwater image processing an important research field, and plenty of CNN (convolutional neural network)-based underwater image enhancement methods have emerged over time. However, the feature-learning ability of existing CNN-based underwater image enhancement is limited. The networks were designed to be complicated or embed other algorithms for better results, which cannot simultaneously meet the requirements of suitable underwater image enhancement effects and real-time performance. Although the composite backbone network (CBNet) was introduced in underwater image enhancement, we proposed OECBNet (optimal underwater image-enhancing composite backbone network) to obtain a better enhancement effect and shorten the running time. Herein, a comprehensive study of different composite architectures in an underwater image enhancement network was carried out by comparing the number of backbones, connection strategies, pruning strategies for composite backbones, and auxiliary losses. Then, a CBNet with optimal performance was obtained. Finally, cross-sectional research of the obtained network with the state-of-the-art underwater enhancement network was performed. The experiments showed that our optimized composite backbone network achieved better-enhanced images than those of existing CNN-based methods.

Keywords: underwater image enhancement; deep learning; composite backbone; composite architectures

1. Introduction

Exploration of the ocean world has attracted more attention in recent years [1]. As a vital part of image processing for revealing and recognition underwater scenes, underwater image enhancement plays an import role in marine resources and marine military fields [2]. Unfortunately, because of differences in illumination conditions in the complicated underwater environments, the captured images are usually of low quality, with characteristics such as having inauthentic color, blurring, and noise, making the enhancement of such images difficult [3]. In order to accomplish this challenging task, many algorithms have been introduced.

In the early stages, the technical breakthroughs in image enhancement often relied on hardware performances [4]. However, advances in image enhancement not only abated the equipment costs but also indicated being more beneficial for accommodating complicated underwater environments. Researchers used mathematical and statistical analyses to modify the channel values of underwater images to obtain images with high visual quality [5]. Sometimes the enhancement effect for underwater images was regarded as an optics problem, and the corresponding algorithms were established on physical models, such as the dark channel prior (DCP) mechanism [6] and the Retinex model [7]. Some

researchers regard underwater image enhancement as a bionic problem [8]. More recently, deep learning gradually became the main underwater image enhancement technology due to its strong modeling ability and efficient inference. Many studies have shown that neural networks obtained enhanced images with higher quality compared to images enhanced by traditional methods [9]. The convolutional neural network (CNN) [10] and generative adversarial network (GAN) [11] were proposed for enhancing underwater images. However, these algorithms are unsuitable because they consume time and space, making them inapplicable.

CBNet shows better performance [12] compared to single-chain networks and provides a broad space for improvement in the network architecture [13]. So herein, CBNet was introduced in the enhancement of underwater images. In order to minimize time and space consumption and meet the actual requirement of marine engineering, CNN models without other embedded algorithms were established. To explore the architectural features of CBNet, comprehensive enhancement research using it in different variants was carried out. An **O**ptimal underwater image **E**nhancement **C**omposite **B**ackbone **Net**work (OECBNet) was finally obtained, which we compared with several state-of-the-art CNN-based methods. Both full-reference indexes and non-reference indexes indicate that the enhancement results of our proposed network are better than recent state-of-the-art CNN-based underwater image enhancement methods. Our proposed network also performs well in real-time image enhancement, with one 350×350 underwater image taking only 2.3×10^{-3} s to enhance.

The strategy of our research is as follows: Section 1 is the introduction to our research. Section 2 describes the related work, including underwater image enhancement methods, underwater image datasets, and CNN backbone networks. Section 3 describes some architectural details of our backbone network. Section 4 makes a comprehensive exploration of the architectural features of CBNet, and the optimal underwater image-enhancing composite backbone network is obtained from this section. Section 5 evaluates our network performance according to cross-sectional research with other methods. Section 6 makes a conclusion of this paper. Contributions in this paper are summarized as shown in the following three points:

- We introduced a composite backbone into CNN-based underwater image enhancement. The composite backbone networks consist of several uniform end-to-end CNN backbones in specific connection strategies. Composite backbone networks were applied throughout the enhancement networks;
- We conducted a comprehensive study about underwater image enhancement using CBNet in different variants by first investigating the impact of backbone number on the enhancement results. Then, the different connection strategies between backbones and some corresponding pruning strategies were introduced in our research. Moreover, the experiments were conducted in auxiliary losses;
- We proposed an optimal underwater image-enhancing composite backbone network. The optimized composite backbone network consists of two backbones in full-connected composition (FCC), and two stages are pruned in the lead backbone. The network accurately expresses objects as well as background colors and is well adapted to the engineering field.

2. Related Work

2.1. Deep Learning-Based Underwater Image Enhancement

Underwater image enhancement methods based on deep learning are classified into five classes according to their network architecture [14]. One idea for constructing image enhancement networks is inspired by semantic segmentation, such as U-Net [15]. Yan et al. [16] proposed a cascaded U-Net for image restoration and achieved enhanced results with less training and inference costs. Moreover, Liu et al. [17] proposed a GAN with an encoder–decoder-based generator, which enabled multi-scale feature extraction and fusion. Other kinds of widely used networks are constructed with multiple branches. Xue et al. [18] enhanced images by predicting the coarse result, veil map, and compensation

map according to a multi-branch network. Li et al. [19] introduced white balance (WB), gamma correction (GC), and histogram equalization (HE)-enhanced images in multiple-branch enhancement and conducted a gated fusion. Similarly, Wu et al. [20] proposed a GAN for enhancement and introduced WB, HE, and DCP enhanced images for multi-branch training in the generator. Wang et al. [21] introduced HSV color space for image adjustment; on this basis, Chen et al. [22] conducted an enhancement fusion of RGB, HSV, and Lab color spaces. Some enhancement methods embedded several units for better feature learning in the networks. Li et al. [9] constructed an E-unit for underwater enhancement, and Chen et al. [23] applied the E-unit in a multi-branch network. Further, depth maps or transmission maps are integrated into some networks [24]. However, to implement these networks, special hardware such as deep cameras are required. Based on the excellent performance of GAN, many researchers tend to upgrade their methods by improving the GAN-based algorithm [25] by incorporating multiple generators [26] or multiple discriminators [27] to enhance underwater images.

2.2. Underwater Dataset

The dataset is a sensitive part when it comes to deep learning. Researchers use different equipment to try and collect images in a variety of marine environments. Liu et al. proposed a real-world underwater image enhancement dataset (RUIE) [28] with a large number of diverse light-scattering images and rich detection targets. Karen et al. created an underwater object tracking dataset (UOT100) [29] that includes 104 underwater videos for object detection. However, the lack of reference images was always a challenge for these datasets to achieve enhancement. To solve the problem, Anwar conducted research on underwater image synthesis and proposed a method for simulating different underwater scenes [9], which can be applied as full-reference enhancement result evaluations. While retaining the real underwater environment, Li et al. proposed an underwater image enhancement benchmark (UIEB) dataset [19] consisting of a reference subset and a challenging subset. The reference dataset included 890 underwater images as well as corresponding reference images, which partially solved the challenge of the lack of reference images and promoted innovations in underwater image enhancement methods.
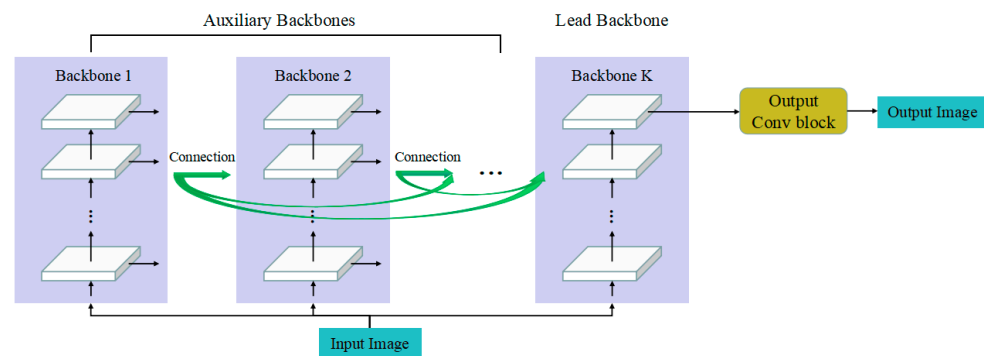
2.3. Backbone Network

The backbone network plays a significant part in CNN models. Earlier networks were often designed to be as deep as possible [30] and connections as dense as possible [31], but their performance generally degenerated with increasing depth. To improve their quality, researchers added branches in different scales into the backbone networks. ResNet, ResNeXt, and Res2Net were proposed in succession [32]. In addition, Liang et al. proposed a recurrent convolutional neural network (RCNN) [33] which recurrently connected each convolution layer, and the contextual information was efficiently integrated. Meanwhile, Zhang et al. [34] proposed a lightweight backbone network called MobileNet, which greatly reduced the time complexity of the backbone network. In order to expand the sensory field of each pixel, many backbone networks embed multi-scale architectures [35], especially feature pyramid blocks [36]. Another inspiration was from U-Net [15], which introduced downsampling and upsampling in backbone networks [37]. Liu et al. proposed a composite backbone network (CBNet) [12], which brought a breakthrough in object detection. Thereafter, they continued to explore the architecture of CBNet and proposed a general framework with novel connection strategies (CBNetV2) [13]. CBNet has now been widely applied to counter the challenges of object detection and semantic segmentation [38]. Zhao et al. improved adjacent higher-level composition in CBNet, achieving the detection and recognition of images in low-quality underwater videos, which brought CBNet into underwater vision [39]. CBNet has been proven to be excellent for feature extraction and is gradually being applied in underwater image enhancement.

## 3. Backbone Network Details

3.1. Global Architecture

The global architecture of CBNet in general is shown in Figure 1. K identical backbones in parallel and the Backbone i are compositely connected with backbones from Backbone 1 to Backbone $i-1$. Output images from the lead backbone are sent to output layers, and output images are obtained from the output convolution block. In addition, each stage obtains images of the same height and weight as input images for pixel-level enhancement. The channel numbers of output images from the stages are constantly the same.

Figure 1. General architecture of CBNet for image enhancement.

3.2. Backbone Architecture

The backbone architecture of our proposed network is as shown in Figure 2, which skillfully ensures that output images are the same height and weight as input images without the shape of input images getting limited. The backbone is empirically constructed with two previous stages, three middle stages, and one last stage in series. Each stage contains a convolutional layer followed by a batch norm layer and a ReLU layer. During the previous stages, the convolutional layers are operated with a 7×7 convolution kernel and padding = 3, and the large convolution kernels expand the receptive field for each pixel, thus improving the global image enhancement effect. During the middle stages, the convolutional layers are operated with a 5×5 convolution kernel and padding = 2, and the introduction of more middle layers for convolution processing improves the learning ability of the network. During the last stages, the convolutional layers are operated with a 3×3 convolution kernel and padding = 1, and last stages are placed before output to accurately enhance the pixelwise detail.
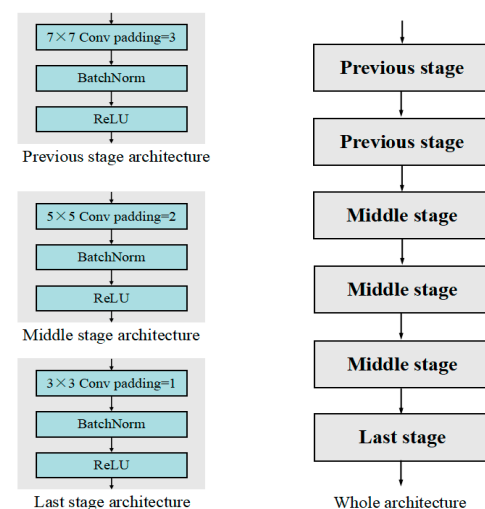
Figure 2. Architecture details of backbones.

3.3. Output Convolution Block

Output blocks play an important role in CNN-based underwater image enhancement for information fusion and channel values' normalization. However, the importance of output blocks was generally ignored in previous underwater image enhancement networks. Our proposed output convolution block incorporated a 3×3 convolutional layer with padding = 1, followed by a Sigmoid layer instead of the traditional ReLU layer. Application of the Sigmoid layer ensures that the pixel channel values of output images lie within the standard range of [0, 1].

4. Research on CBNet-Based Underwater Image Enhancement

In this section, we explore the architectural features of CBNet. First, we introduce some experiment details. In Section 4.2, we analyze the impact of backbone number on underwater image enhancement. In Section 4.3, we evaluate the underwater image enhancement performance of different variants. In Section 4.4, we implement pruning strategies in some variants. In Section 4.5, we construct auxiliary losses for partial special variants. Finally, we make a summary of composite backbone architecture and choose a network as the optimal composite backbone network for subsequent comparison.

4.1. Implementation Details

We applied the UIEB dataset [22] to our study. To construct a training set, we randomly chose 800 underwater images as well as their corresponding reference images from a reference subset. Other images in the UIEB dataset were treated as test images. All images were preprocessed by stretching to 350×350, and training set images were cut into 320×320 at random. Our loss function was a weighted sum of $L1$ loss and SSIM loss as follows:

$$L(I_P, I_R) = \omega_{l_1} L_{l_1}(I_P, I_R) + \omega_{\text{SSIM}} L_{\text{SSIM}}(I_P, I_R) \tag{1}$$

where ω_{l_1} and ω_{SSIM} are weight parameters used to balance two loss components. Weight parameters are selected as ω_{l_1} = 2, ω_{SSIM} = 1. The two loss components L_{l_1} and L_{SSIM} are calculated as follows:

$$L_{l_1}(I_P, I_R) = \frac{1}{C^{I_R} H^{I_R} W^{I_R}} \sum \left| t^{I_P}(i) - t^{I_R}(i) \right| \tag{2}$$

$$L_{\text{SSIM}}(I_P, I_R) = 1 - \frac{1}{C^{I_R} H^{I_R} W^{I_R}} \sum \frac{2\mu_{I_P}(i)\mu_{I_R}(i) + C_1}{\mu_{I_P}^2(i) + \mu_{I_R}^2(i) + C_1} \times \frac{2\sigma_{I_P I_R}(i) + C_2}{\sigma_{I_P}^2(i) + \sigma_{I_R}^2(i) + C_2} \tag{3}$$

where I_P represents the predicted image and I_R represents the reference image. $t^{I_P}(i)$ and $t^{I_R}(i)$ are the corresponding channel values in a certain pixel i. $\mu_{I_P}(i)$ and $\mu_{I_R}(i)$ are average channel values of a 11^2 square interval with center i. $\sigma_{I_P}(i)$ and $\sigma_{I_R}(i)$ represent corresponding standard deviations. $\sigma_{I_P I_R}(i)$ corresponds to the covariance, while parameters $C_1 = 0.02$ and $C_2 = 0.03$ are selected to stabilize the SSIM function. C, H, and W mean channels, height, and weight, respectively.

We applied ADAM optimizer for network training and initialized the learning rate as 1×10^{-3} and decreased it by 1×10^{-5} in each epoch. We extracted the trained network models after 100 epochs, which means it takes nearly two hours to train each model. Pytorch as well as Anaconda were jointly used with Nvidia GTX 3080 GPU for programming.

We used the rest of the 90 pairs of images from the UIEB reference subset for network evaluation, and three classical full-reference indications, namely: mean square error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM), were adopted. A lower MSE or higher PSNR score indicated that the pixel values of the predicted results were closer to references. A higher SSIM indicated that the predicted result was more similar to the reference. However, these indications could only evaluate images from a single channel. To make evaluation results more rigorous, we conducted evaluations from both gray-scale images and RGB images as follows:

$$E_{\text{RGB}}(I, I) = (E(I_R, I'_R) + E(I_G, I'_G) + E(I_B, I'_B))/3 \tag{4}$$

$$E_{\text{gray}}(I, I') = E(\text{RGB2 gray}(I), \text{RGB2 gray}(I')) \tag{5}$$

where E represents evaluation functions of indications such as MSE, PSNR, or SSIM. I is the enhanced image while I_R, I_G, and I_B are corresponding maps in R, G, and B channels; I' is the reference image while I'_R, I'_G, and I'_B are corresponding maps in R, G, and B channels.
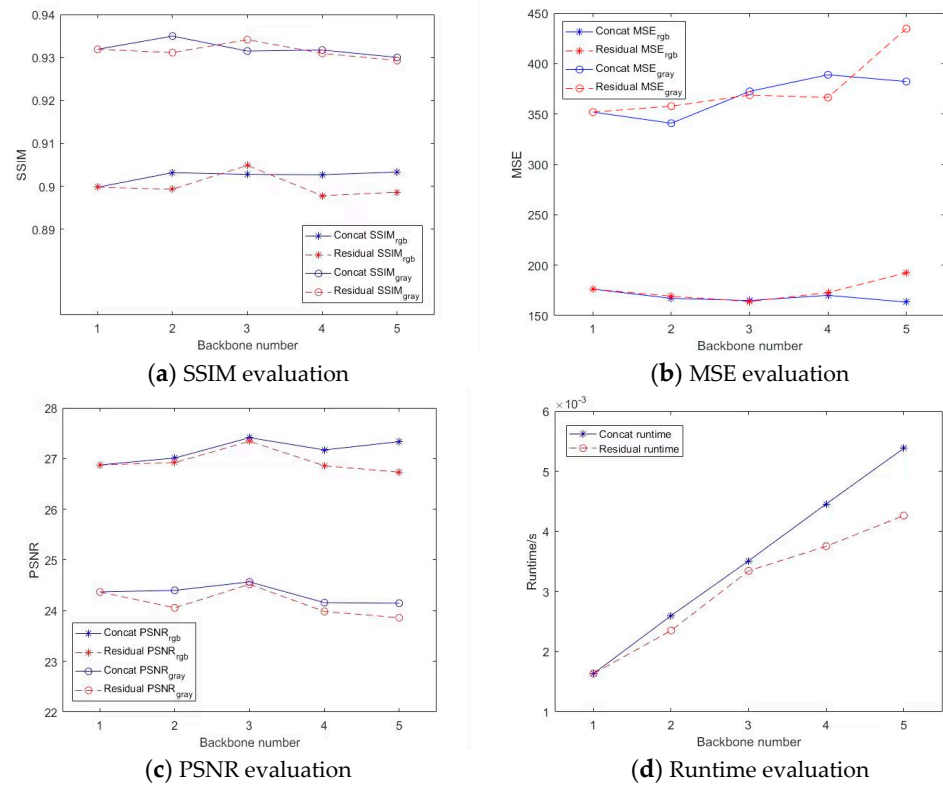
4.2. Backbone Number Analysis

CBNet was a combination of K backbones with the same architecture. We define the backbones as Bone 1, Bone 2, Bone 3, ..., Bone K. For any defined backbone, if $k = 1, 2, \ldots, K - 1$, Bone k is considered an auxiliary backbone; if $k = K$, Bone k is the lead backbone. Generally, there are two approaches for information fusion between backbones, that is, residual fusion and concat fusion. Concat fusion is to amalgamate several images with the same shape in the specified dimension, and the channel number of the output image is the sum of input image channel numbers. Residual fusion is a simple image subtraction. The traditional same-level composition (SLC), as shown in Figure 4a, is applied in backbone number analysis; thus, the lth stage transformation of concat-fused information and residual-fused information are shown respectively:

$$x_k^l = F_C^l(Cat(x_1^{l-1} + x_2^{l-1} + \cdots + x_k^{l-1})), \ 2 \le l \le 7, \ 2 \le k \le K \tag{6}$$

$$x_k^l = F_R^l(x_k^{l-1} - x_{k-1}^{l-1}), \ 2 \le l \le 7, \ 2 \le k \le K \tag{7}$$

where Cat represents the concat fusion of images. F_C^l is the combinational function of concat-fused information in the lth stage, F_R^l is the combinational function of residual-fused information in the lth stage, while x_k^l is the output image from the lth stage in the kth backbone. When $l = 7$, F_C^7 and F_R^7 indicate the combinational function of the output block, while x_k^7 indicates the output-enhanced image.

We conducted analyses of underwater image enhancement from networks with different backbone numbers. The output channel number from each stage was set as 24 to meet the internal memory demand. Underwater image enhancement analyses were conducted, as shown in Figure 3, details of analyses from different backbone number as shown in Table S1. According to SSIM evaluations in Figure 3a, networks with concat fusion obtained the best SSIM score when $K = 3$, and networks with residual fusion obtained the best SSIM score when $K = 2$. According to MSE evaluations in Figure 3b, as the number of backbones increased, MSE_{gray} indicators increased significantly, while MSE_{RGB} from concat fusion networks obtained a weak downward trend. In our case, networks with concat fusion obtained better MSE scores than networks with residual fusion. The advantage of networks with concat fusion became evident according to RSNR evaluation in Figure 3c, and the best PSNR scores were obtained when $K = 3$. According to Figure 3d, the average runtime of concat fusion networks increased linearly with the number of backbones. Compared with concat fusion networks, the runtime of residual fusion networks decreased to some extent. According to the above results, increasing the number of backbones can increase the amount of information and thus increase the possibility of obtaining high-quality results. However, with the increase in information, the time complexity will increase, and the interference information will increase, too. What is more, the introduction of too-deep network enhancement will gradually distort output images. Thus, excessively increasing the auxiliary backbone is often counterproductive. Suitable backbone architectures would appear when $K = 2$ or 3.

(**a**) SSIM evaluation

(**b**) MSE evaluation

(**c**) PSNR evaluation

(**d**) Runtime evaluation

Figure 3. Analyses from different backbone numbers.

4.3. Connection Strategy Analysis

The analysis in Section 4.2 showed that concat fusion networks achieve relatively suitable underwater image enhancement results when $K = 2$ or 3. To minimize the time and space complexity, we employed CBNet with two concat-fused backbones for further exploration. To further explore the CBNet architecture, we constructed several variants of CBNet in different connection strategies besides the traditional SLC (Figure 4). Details of different composite strategies are explained in Sections 4.3.1–4.3.5. The output channel number from each stage was set as 32.

(a) SLC   (b) AHLC   (c) ALLC   (d) DSLC   (e) DHLC   (f) FCC
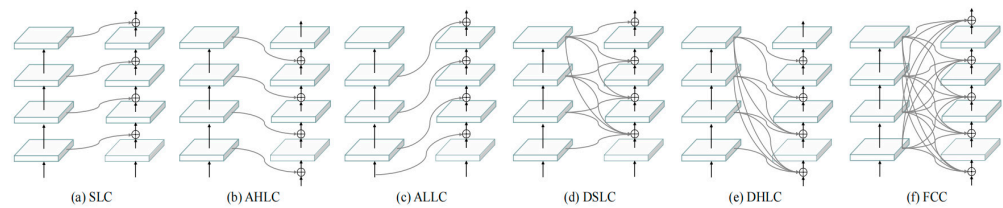
Figure 4. Architecture of different composite strategies for CBNet.

4.3.1. Adjacent Higher-Level Composition (AHLC)

To obtain stable image enhancement results, higher-level features can be taken to enhance the lower-level features. In the AHLC variant shown in Figure 4b, output information of the adjacent higher-level stage from the auxiliary backbone was sent to the lead backbone. The *l*th stage transformation of AHLC is shown in Equation (8):

$$x_2^l = F_C^l(Cat(x_1^l, x_2^{l-1})), 1 \leq l \leq 6 \tag{8}$$

where $l = 1$, x_k^0 indicates input image.

4.3.2. Adjacent Lower-Level Composition (ALLC)

Another line for stabilizing image enhancement results is to take lower-level features to enhance the higher-level features, which is just opposite to AHLC. As shown in Figure 4c, ALLC sent the adjacent lower-level stage output information to the lead backbone. The *l*th stage transformation of ALLC is shown in Equation (9):

$$x_2^l = F_C^l(Cat(x_1^{l-2}, x_2^{l-1})),\ 2 \le l \le 7 \tag{9}$$

4.3.3. Dense Same-Level Composition (DSLC)

As shown in Figure 4d, the output information of each stage from the lead backbone was fused with the output information of non-lower-level stages from the auxiliary backbone and sent to the next stage. The architecture of DSLC is similar to that of DenseNet [33]. The lth stage transformation of DSLC is as shown in Equation (10):

$$x_2^l = F_C^l(Cat(x_1^{l-1}, x_1^l, \cdots, x_1^6, x_2^{l-1})),\ 2 \le l \le 7 \tag{10}$$

4.3.4. Dense Higher-Level Composition (DHLC)

As shown in Figure 4e, the output information of higher-level stages was sent to the lead backbone. The *l*th stage transformation of DHLC is as shown in Equation (11):

$$x_2^l = F_C^l(Cat(x_1^l, x_1^{l+1}, \cdots, x_1^6, x_2^{l-1})),\ 1 \le l \le 6 \tag{11}$$

4.3.5. Full-Connected Composition (FCC)

To connect the auxiliary backbone to the leading backbone as compactly as possible, each auxiliary stage backbone output information was sent to all lead backbone stages, as shown in Figure 4f. The *l*th stage transformation of FCC is as shown in Equation (12):

$$x_2^l = F_C^l(Cat(x_1^1, x_1^2, \cdots, x_1^6, x_2^{l-1})),\ 1 \le l \le 7 \tag{12}$$

We then evaluated underwater image enhancement results from networks with different composite strategies described above. The output channel number from each stage was set as 32. Evaluations of the enhanced underwater images from different connection strategies are shown in Table 1, details as shown in Table S2.

Table 1. Enhancement results evaluation for different connection strategies.

Method	SSIM_{RGB}	$\text{SSIM}_{\text{gray}}$	MSE_{RGB}	MSE_{gray}	PSNR_{RGB}	$\text{PSNR}_{\text{gray}}$
SLC	0.9025	0.9323	172.95	390.66	27.224	24.374
AHLC	0.9017	0.9314	164.52	358.08	27.322	24.523
ALLC	0.9045	0.9344	170.36	369.54	26.983	24.107
DSLC	0.9059	0.9350	164.38	371.45	27.201	24.201
DHLC	0.9024	0.9318	163.32	362.24	27.370	24.445
FCC	0.9080	0.9375	173.81	384.59	27.026	24.047

According to the evaluations, FCC achieved the best scores in SSIM indications, DHLC achieved the best scores in MSE and PSNR indications, and AHLC achieved the best scores in MSE_{gray} and $\text{PSNR}_{\text{gray}}$. The results revealed that FCC obtained enhancement images almost similar to reference images, DHLC obtained enhancement images closely resembling reference images, while AHLC obtained enhancement images closer to reference images from gray-scale than other connection strategies.

Further, Table 2 shows the runtime comparison among networks with different connection strategies. According to the results, due to the reduction in the connection between the lead backbone and the auxiliary backbone, ALLC completed underwater image enhancement within the shortest runtime (2.55×10^{-3} s per image on average). Meanwhile, DHLC constructed pretty complex connections between the lead backbone and the auxil-

iary backbone, thus completing underwater image enhancement in the longest runtime. Notably, FCC constructed similar concat connections before each stage, which simplified the network procedure; thus, FCC also enhanced images relatively fast, with an average runtime of only 2.59×10^{-3} s.

Table 2. Runtime comparison for networks with different connection strategies.

Method	Average Runtime/s
SLC	0.00261
AHLC	0.04381
ALLC	0.00255
DSLC	0.13251
DHLC	0.13596
FCC	0.00259

4.4. Pruning Strategy Analysis

To reduce the excessive connection between network backbones, we simplified the architecture of CBNet to some extent by pruning. Pruning strategies were applied in CBNet when $K = 2$. As shown in Figure 5, p_i represents pruning i stages from the lead backbone; two backbones share the first i stages. The $i+1$th stage transformation is as shown in Equation (13):

$$x_2^{i+1} = F_C^{i+1}(Cat(Cat(x_1), x_1^i)), \ 0 \leq i \leq 6 \tag{13}$$

where $Cat(x_1)$ represents the connection from the auxiliary backbone to the lead backbone. When $i = 0$, the network is not pruned; when $i = 6$, the network is a single backbone network. According to Section 4.3, enhancement results from FCC and DHLC achieved excellent scores, and both FCC and DHLC variants were applied for subsequent studies.
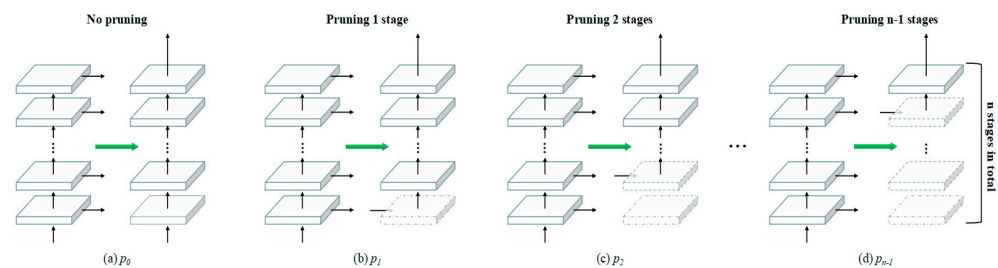
Figure 5. Architecture of pruning strategies for CBNet.

We conducted analyses of underwater image enhancement from networks with different pruning strategies, as shown in Figure 6, details of analyses from different pruning strategies as shown in Table S3. According to the SSIM evaluation in Figure 6a, FCC obtained better SSIM scores than DHLC. Both FCC and DHLC obtained the best enhancement result when pruning two stages. Moreover, FCC(p_2) achieved $\text{SSIM}_{\text{gray}} = 0.9409$ and $\text{SSIM}_{\text{RGB}} = 0.9119$, which is a breakthrough in full-reference underwater image enhancement. According to the MSE evaluation in Figure 6b, the best MSE scores for FCC and DHLC were also obtained when pruning two stages, with MSE scores for FCC being slightly lower than those for DHLC. According to Figure 6c, the best PSNR scores for DHLC were obtained in p_2. On the contrary, the best PSNR scores for FCC were obtained in p_3 and p_4, respectively. According to a runtime evaluation in Figure 6d, the runtime of FCC had an inverse relationship with pruning stages, with runtime decreasing smoothly with an increase in pruning stages. The runtime of FCC remained in the range of $1.4 \times 10^{-3} \sim 2.6 \times 10^{-3}$ s. The runtime of DHLC from p_0 to p_3 was excessively long, and there was a sudden decline in runtime between p_3 and p_4. In conclusion, FCC(p_2) obtained the best underwater image enhancement results, and DHLC(p_2)-enhanced results obtained the best PSNR scores. These architectures were applied in subsequent studies.

(**a**) SSIM evaluation

(**b**) MSE evaluation

(**c**) PSNR evaluation

(**d**) Runtime evaluation
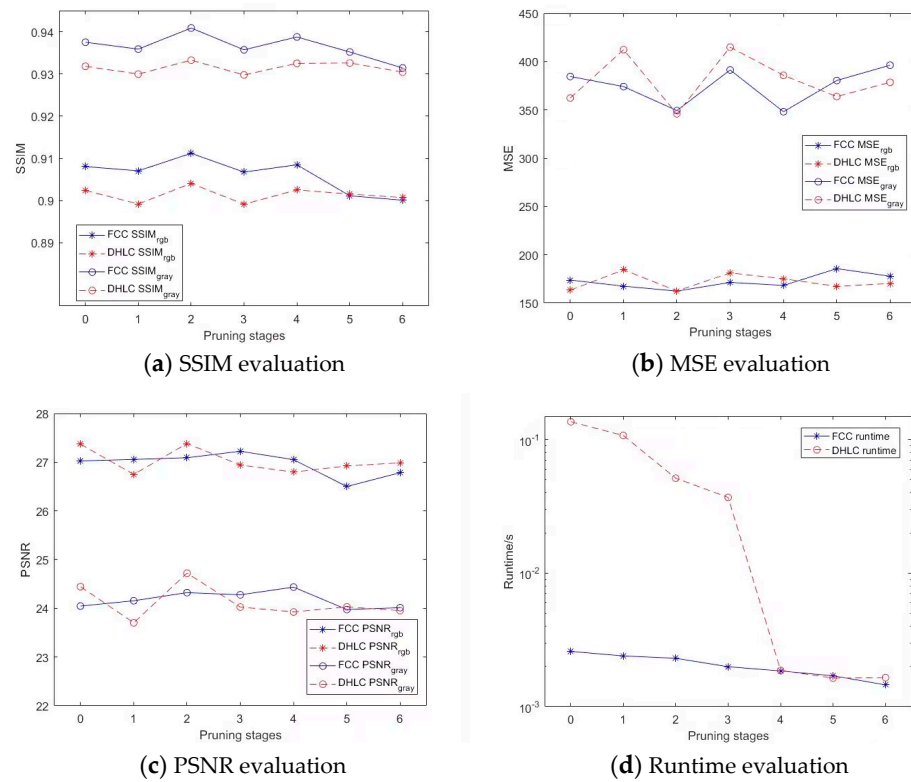
Figure 6. Analyses from pruning strategies.

4.5. Auxiliary Loss Analysis

Considering the special architecture of AHLC and DHLC, output information of the lead backbone's last stage was directly transmitted to Output Conv Block. As a result, the input information of Output Conv Block had the same channel number as the output information of the auxiliary backbones. Thus, auxiliary losses could be introduced in the above variants, and information from the auxiliary backbones was sent to the Output Conv Block, generating an auxiliary image. The auxiliary losses were computed as shown in Equation (14):

$$L(I_{aux}^k, I_R) = \omega_{l_1} L_{l_1}(I_{aux}^k, I_R) + \omega_{\text{SSIM}} L_{\text{SSIM}}(I_{aux}^k, I_R) \tag{14}$$

where I_{aux}^k represents the auxiliary image from the kth backbone, $k = 1, 2, \ldots, k - 1$. The total loss was constructed as a weighted sum of the main loss as well as auxiliary losses, as shown in Equation (15):

$$L_{total} = L(I_P, I_R) + \sum_{k=1}^{K-1} \lambda_k L(I_{aux}^k, I_R) \tag{15}$$

where λ_k is weight for auxiliary loss from I_{aux}^k.

We introduced auxiliary losses in AHLC, DHLC as well as DHLC(p_2) according to Sections 4.3 and 4.4. The architecture is shown in Figure 7. Moreover, we put two auxiliary strategies into the experiment; one is the full-auxiliary strategy (written down as *f-a*), where loss weight is selected as a constant that $\lambda_1 = 0.3$, and the other is the semi-auxiliary strategy (written down as *s-a*) where during first 50 epochs, $\lambda_1 = 0.3$ and $\lambda_1 = 0$ later on.
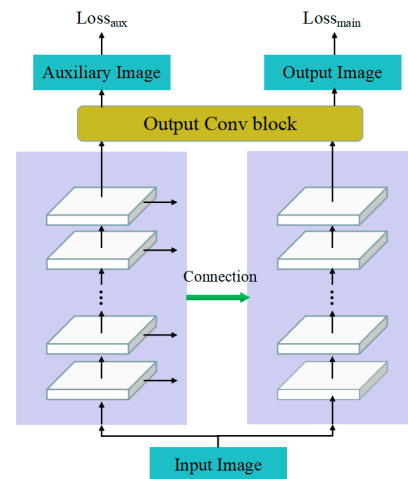
**Figure 7.** Architecture of CBNet ($K = 2$) with auxiliary loss.

We evaluated semi-auxiliary and full-auxiliary strategies underwater image enhancement results, then compared them with enhancement results without auxiliary losses. Evaluations of underwater image enhancement results are shown in Figures 8–10, details as shown in Table S4.

(**a**) SSIM$_{\text{RGB}}$

(**b**) SSIM$_{\text{gray}}$
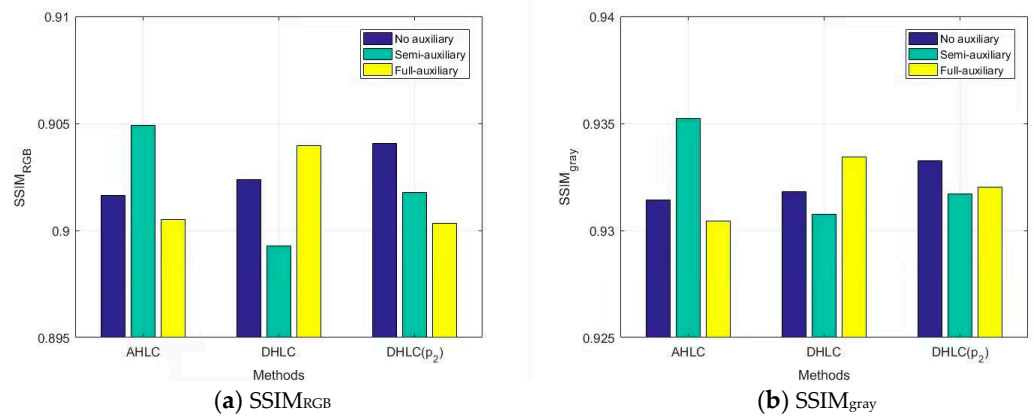
Figure 8. SSIM evaluation on auxiliary loss analysis.

(**a**) MSE$_{\text{RGB}}$

(**b**) MSE$_{\text{gray}}$
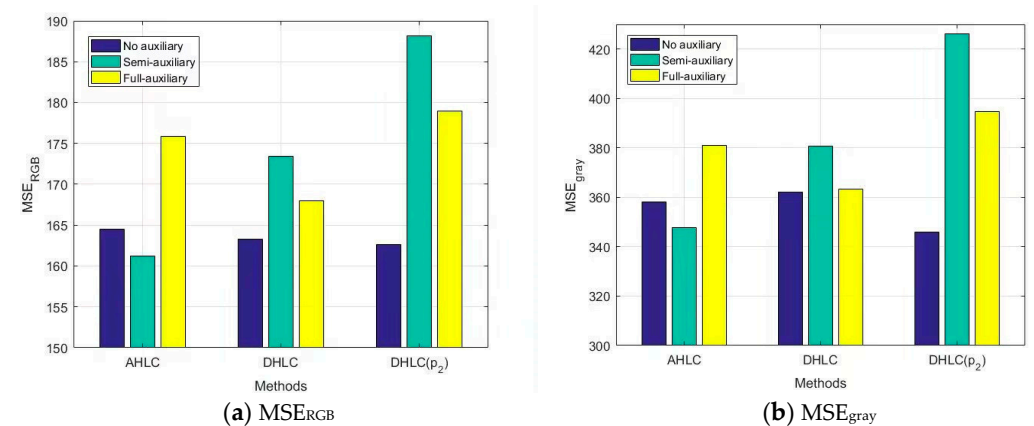
Figure 9. MSE evaluation on auxiliary loss analysis.

(**a**) PSNR$_{RGB}$ (**b**) PSNR$_{gray}$
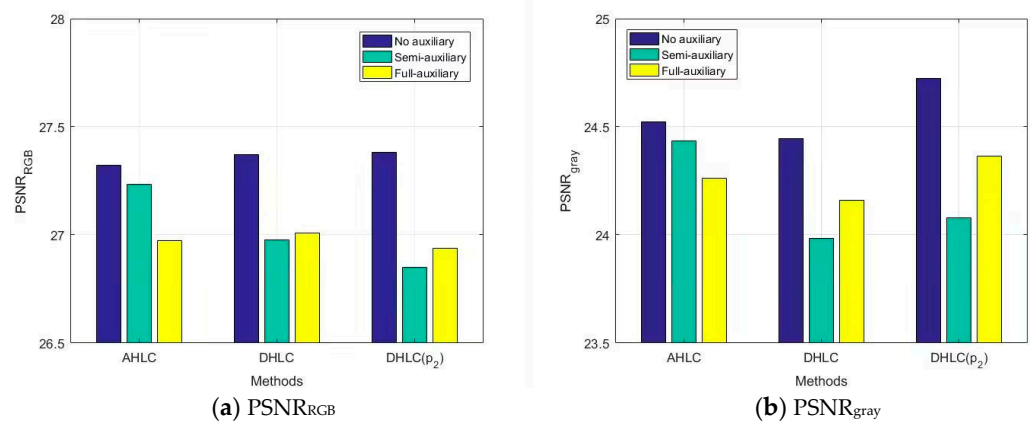
Figure 10. PSNR evaluation on auxiliary loss analysis.

According to SSIM evaluation (Figure 8) and MSE evaluation (Figure 9), AHLC obtained the best scores in the semi-auxiliary strategy (MSE$_{RGB}$ = 161.22), which is also a breakthrough in MSE evaluation. For DHLC, the full-auxiliary strategy and the method without auxiliary losses obtained better results than the semi-auxiliary strategy. DHLC(p_2) obtained the best scores without auxiliary losses. However, according to Figure 10, auxiliary strategies did not perform well in RSNR. As auxiliary losses are abandoned during network testing, the introduction of auxiliary losses does not affect the network runtime.

4.6. Summary

According to the comprehensive analyses above, various network architectures and optimization strategies of CBNet have achieved significant effects in underwater image enhancement. Among them, FCC with two pruned stages (FCC(p_2)), DHLC with two pruned stages (DHLC(p_2)), and AHLC with the semi-auxiliary strategy (AHLC(*s-a*)) obtained excellent image enhancement results. We conducted a comparison among the three networks, as shown in Table 3, where FCC(p_2) obtained the best scores in SSIM evaluations, DHLC(p_2) obtained the best scores in PSNR evaluations, and AHLC(*s-a*) obtained the best scores in MSE$_{RGB}$. Considering FCC(p_2) can obtain effective underwater image enhancement in the shortest time, hence meeting the demand of real-time performance, FCC(p_2) was selected as the optimal underwater image-enhancing composite backbone network (OECBNet).

Table 3. Comparison of evaluation results for three excellent networks.

Method	SSIM$_{RGB}$	SSIM$_{gray}$	MSE$_{RGB}$	MSE$_{gray}$	PSNR$_{RGB}$	PSNR$_{gray}$	Runtime
FCC(p_2)	0.911	0.941	162.4	349.3	27.09	24.32	0.0023
DHLC(p_2)	0.904	0.933	162.6	345.9	27.38	24.72	0.0514
AHLC(*s-a*)	0.905	0.935	161.2	347.7	27.23	24.44	0.0438

5. Cross-Sectional Research

To verify the performance of our proposed OECBNet, we compared OECBNet with several state-of-the-art CNN-based methods, including UWCNN [9], Water-Net [19], UIEC2-Net [21], IW-Net [23], and MC-CBNet [22]. Besides the 90 pairs of images from the UIEB reference subset, 60 images from the UIEB challenge subset were applied to the analysis. Both full-reference comparison and non-reference comparison were applied in the cross-sectional study. We applied two indicators in the non-reference comparison, one being underwater color image quality evaluation (*UCIQE*) [40], which consists of color density σ_c, saturation μ_s, and contrast con_l, as shown in Equation (16),

$$UCIQE = 0.4680\sigma_c + 0.2745con_l + 0.2576\mu_s, \qquad (16)$$

and the other being underwater image quality measure (*UIQM*) [41], which is a combination of three sub-indicators: underwater image colorfulness measurement (*UICM*), underwater image sharpness measurement (*UISM*), and underwater image contrast measurement (*UIConM*). The expression is shown in Equation (17):

$$UIQM = 0.0282UICM + 0.2953UISM + 3.5753UIConM \tag{17}$$

5.1. Reference Subset Evaluation

First, we tested our network as well as state-of-the-art CNN-based methods in the reference subset, and some underwater image enhancement results are shown in Figure 11.
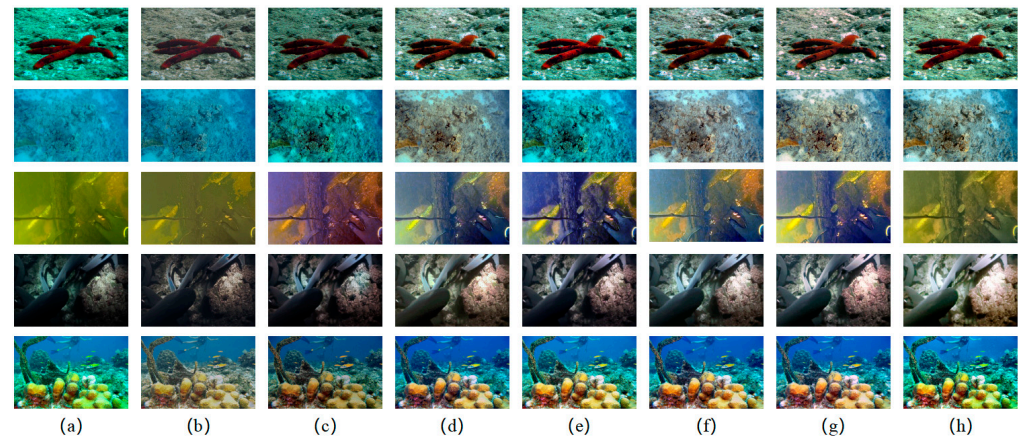
Figure 11. Subjective comparison of underwater image enhancement results using a reference subset. (**a**) Raw images. (**b**) Results from UWCNN. (**c**) Results from Water-Net. (**d**) Results from UIEC2-Net. (**e**) Results from IW-Net. (**f**) Results from MC-CBNet. (**g**) Results from the proposed OECBNet. (**h**) Reference images.

As they are affected by natural illumination and differences in water quality, underwater images bear different features. In shallow water, the color features of images are reserved greatly due to the relatively high illumination. As the light gradually spreads into the deep underwater, the red, green, and blue spectra disappear in succession; hence, many underwater images tend to have cool tones, such as green and blue. On the other hand, images from the water–sediment interphase tend to be yellowish-brown, while images from low-illumination areas appear dim. According to the subjective comparison shown in Figure 11, it is visible that our proposed OECBNet obtained images the closest to reference images from high illumination, low illumination, and greenish or blueish underwater images and obtained clearer images from yellowish underwater images even than reference images. This indicated that OECBNet effectively eliminated the blue and green color bias in underwater images, reduced the effect of sediment on underwater images, and accurately expressed the color of objects and the environment. In addition, our proposed method retained detailed features of the underwater images and enhanced them to appear brighter and clearer.

A full-reference comparison of OECBNet as well as state-of-the-art CNN-based methods is shown in Table 4. OECBNet obtained the best scores in SSIM and MSE evaluations but did not achieve the best PSNR scores, indicating that images obtained from OECBNet still bear great deviations in some pixels.

Table 4. Full-reference comparison of enhancement results in reference subset.

Method	SSIM$_{RGB}$	SSIM$_{gray}$	MSE$_{RGB}$	MSE$_{gray}$	PSNR$_{RGB}$	PSNR$_{gray}$
UWCNN	0.8443	0.8799	496.32	1191.6	22.550	19.470
Water-Net	0.8840	0.9188	380.14	874.80	23.738	20.946
UIEC^{2}Net	0.9034	0.9336	168.35	368.49	27.039	24.238
IW-Net	0.9062	0.9348	223.10	473.09	27.098	24.496
MC-CBNet	0.9048	0.9349	168.11	385.16	27.375	24.553
OECBNet	0.9112	0.9409	162.40	349.29	27.091	24.322

A non-reference comparison of OECBNet and the state-of-the-art CNN-based methods is shown in Table 5. OECBNet obtained the best scores in *UCIQE*, *UIQM*, *UICM*, and *UISM* and had better scores than reference images in *UCIQE*, *UIQM*, *UISM*, and *UIConM*, which verified the excellent performance of our proposed network.

Table 5. Non-reference comparison of enhancement results in reference subset.

Method	*UCIQE*	*UIQM*	*UICM*	*UISM*	*UIConM*
Raw	0.4525	1.9821	4.9778	3.5095	0.2252
UWCNN	0.4416	2.5893	4.0856	4.6135	0.3109
Water-Net	0.5341	2.5043	5.4714	4.5885	0.2783
UIEC^{2}Net	0.5288	2.6618	6.5597	4.8831	0.2895
IW-Net	0.5559	2.5874	6.8383	4.7999	0.2733
MC-CBNet	0.5520	2.6580	6.9278	4.8821	0.2856
OECBNet	0.5612	2.7178	7.1059	5.0357	0.2882
Reference	**0.5595**	**2.5024**	**7.5522**	**4.6337**	**0.2576**

The runtime analysis is shown in Table 6. Our proposed network had an obvious real-time advantage compared with most underwater image enhancement networks and achieved underwater image enhancement much swifter than state-of-the-art CNN-based methods, except for UWCNN. It can be seen that the optimization of the architecture has achieved an obvious effect on time complexity and effectively solved the problem of the runaway growth of the time complexity caused by the complex network.

Table 6. Runtime analysis of enhancement results.

Method	Average Runtime/s
UWCNN	0.00143
Water-Net	0.23477
UIEC^{2}Net	0.11010
IW-Net	0.31910
MC-CBNet	0.37713
OECBNet	0.00230

5.2. Challenging Subset Evaluation

To further verify the excellent image enhancement capability of our network, we tested our network against the state-of-the-art CNN-based methods using a challenging subset from the UIEB dataset. The results are shown in Figure 12.
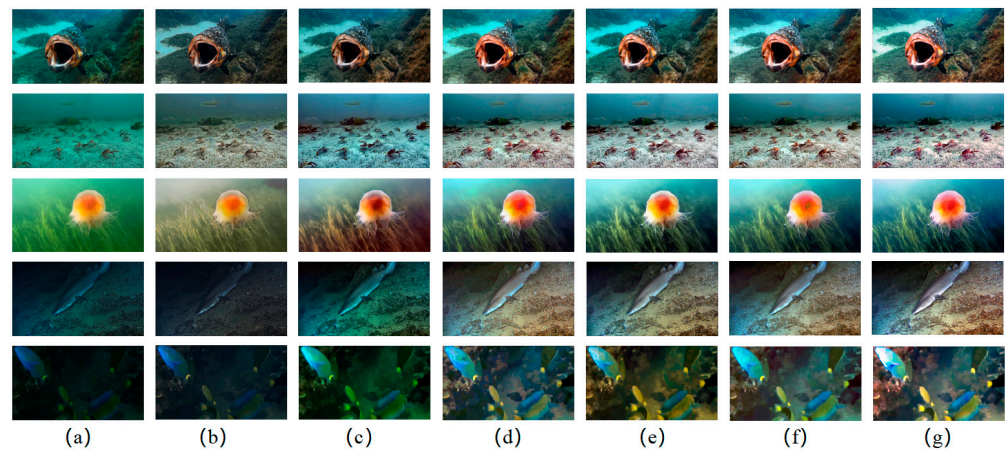
Figure 12. Subjective comparison of underwater image enhancement results using a challenging subset. (**a**) Raw images. (**b**) Results from UWCNN. (**c**) Results from Water-Net. (**d**) Results from UIEC2-Net. (**e**) Results from IW-Net. (**f**) Results from MC-CBNet. (**g**) Results from the proposed OECBNet.

Our proposed method eliminated color cast and obtained more realistic images than other strategies. For underwater images from low-visibility environments, OECBNet clearly restored the underwater scene. According to the challenge subset, the texture structures of the images are relatively simple, and the visibility is relatively low. It is relatively harder to see differences among the enhanced images from different algorithms on the macro level. Although there were no significant differences among IW-Net (e), MC-CBNet (f), and OECBNet (g) enhanced results, our proposed OECBNet achieved better color control in detail, which can be clearly represented by quantitative evaluation.

Due to the lack of reference images in the challenging subset, we applied non-reference indicators to quantitatively evaluate the enhanced underwater images. The comparison of OECBNet with state-of-the-art CNN-based methods is shown in Table 7. OECBNet-enhanced images obtained the best scores in *UCIQE*, *UIQM*, *UICM*, and *UISM*, hence achieving the best underwater image enhancement results.

Table 7. Quantitative evaluation of enhancement results in the challenging subset.

Method	UCIQE	UIQM	UICM	UISM	UIConM
Raw	0.3907	1.6933	3.1112	2.4060	0.2504
UWCNN	0.3877	2.1985	2.5470	3.3318	0.3196
Water-Net	0.5132	2.1455	4.4568	3.2896	0.2932
UIEC^{2}Net	0.5056	2.3049	4.7146	3.7595	0.2970
IW-Net	0.5369	2.2320	4.7800	3.4878	0.2985
MC-CBNet	0.5302	2.2942	5.0663	3.7356	0.2932
OECBNet	0.5503	2.3240	5.3689	3.7981	0.2940

6. Conclusions

In this paper, we applied a composite backbone network (CBNet) in underwater image enhancement by first constructing the global architecture of the proposed CBNet as well as backbone details, then conducting comprehensive research on the architecture of CBNet. Since the backbone is an important part of CBNet, we analyzed the impact of backbone number for CBNet-based underwater image enhancement, then compared the performances of different connection strategies in CBNet, which includes same-level composition (SLC), adjacent higher-level composition (AHLC), adjacent lower-level composition (ALLC), dense same-level composition (DSLC), dense higher-level composition (DHLC), and full-connected composition (FCC). Due to the suitable performance of DHLC and FCC, both of them were selected for pruning strategy analysis. In addition, the auxiliary loss was also applied in CBNet research. To summarize, we selected a network

with the best enhancement results as the optimal underwater image-enhancing composite backbone network (OECBNet) for subsequent research. According to the cross-sectional research conducted, the images obtained by our proposed network accurately expressed the colors of objects as well as the colors of corresponding environments while retaining fine feature details. According to quantitative analyses, our proposed network obtained better scores than state-of-the-art CNN-based methods in both full-reference indications and non-reference indications. In addition, our proposed network also performed well in real-time image enhancement and achieved a record 2.3×10^{-3} s per image. We postulate that CBNet, with more than three backbones and multiple connection strategies, might achieve better performance in underwater image enhancement and recommend further research into these architectures. We expect that our proposed underwater image enhancement network will be applied in ROV for taking underwater videos and underwater detection.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/biomimetics8030275/s1, Table S1: Analyses from different backbone number; Table S2: Analyses from different connection strategies; Table S3: Analyses from different pruning strategies; Table S4: Auxiliary loss analysis.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to it being only available to teams interested in collaboration.

Conflicts of Interest: The authors declare that they have no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Cong, Y.; Fan, B.; Hou, D.; Fan, H.; Liu, K.; Luo, J. Novel event analysis for human-machine collaborative underwater exploration. *Pattern Recognit.* **2019**, *96*, 106967. [CrossRef]
2. Zhou, Y.; Li, B.; Wang, J.; Rocco, E.; Meng, Q. Discovering unknowns: Context-enhanced anomaly detection for curiosity-driven autonomous underwater exploration. *Pattern Recognit.* **2022**, *131*, 108860. [CrossRef]
3. Zhang, W.; Zhuang, P.; Sun, H.-H.; Li, G.; Kwong, S.; Li, C. Underwater Image Enhancement via Minimal Color Loss and Locally Adaptive Contrast Enhancement. *IEEE Trans. Image Process.* **2022**, *31*, 3997–4010. [CrossRef] [PubMed]
4. Li, W.; Hu, W.; Dong, T.; Qu, J. Depth Image Enhancement Algorithm Based on RGB Image Fusion. In Proceedings of the 2018 11th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 8–9 December 2018; Volume 2, pp. 111–114. [CrossRef]
5. Ancuti, C.O.; Ancuti, C.; De Vleeschouwer, C.; Bekaert, P. Color Balance and Fusion for Underwater Image Enhancement. *IEEE Trans. Image Process.* **2018**, *27*, 379–393. [CrossRef]
6. Peng, Y.-T.; Cao, K.; Cosman, P.C. Generalization of the Dark Channel Prior for Single Image Restoration. *IEEE Trans. Image Process.* **2018**, *27*, 2856–2868. [CrossRef]

7. Lv, X.; Sun, Y.; Zhang, J.; Jiang, F.; Zhang, S. Low-light image enhancement via deep Retinex decomposition and bilateral learning. *Signal Process. Image Commun.* **2021**, *99*, 116466. [CrossRef]

8. Gao, S.-B.; Zhang, M.; Zhao, Q.; Zhanga, X.-S.; Li, Y.-J. Underwater Image Enhancement Using Adaptive Retinal Mechanisms. *IEEE Trans. Image Process.* **2019**, *28*, 5580–5595. [CrossRef]

9. Li, C.; Anwar, S.; Porikli, F. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognit.* **2020**, *98*, 107038. [CrossRef]

10. Li, C.; Guo, C.; Chen, C.L. Learning to Enhance Low-Light Image via Zero-Reference Deep Curve Estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 4225–4238. [CrossRef]

11. Guo, Y.; Li, H.; Zhuang, P. Underwater Image Enhancement Using a Multiscale Dense Generative Adversarial Network. *IEEE J. Ocean. Eng.* **2020**, *45*, 862–870. [CrossRef]

12. Liu, Y.; Wang, Y.; Wang, S.; Liang, T.; Zhao, Q.; Tang, Z.; Ling, H. CBNet: A Novel Composite Backbone Network Architecture for Object Detection. *Proc Conf AAAI Artif Intell* **2020**, *34*, 11653–11660. [CrossRef]

13. Liang, T.; Chu, X.; Liu, Y.; Wang, Y.; Tang, Z.; Chu, W.; Chen, J.; Ling, H. CBNet: A Composite Backbone Network Architecture for Object Detection. *IEEE Trans. Image Process.* **2022**, *31*, 6893–6906. [CrossRef]

14. Anwar, S.; Li, C. Diving deeper into underwater image enhancement: A survey. *Signal Process. Image Commun.* **2020**, *89*, 115978. [CrossRef]

15. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention 2015*; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241. [CrossRef]

16. Yan, L.; Zhao, M.; Liu, S.; Shi, S.; Chen, J. Cascaded transformer U-net for image restoration. *Signal Process.* **2023**, *206*, 108902. [CrossRef]

17. Liu, X.; Gao, Z.; Chen, B.M. MLFcGAN: Multilevel Feature Fusion-Based Conditional GAN for Underwater Image Color Correction. *IEEE Geosci. Remote. Sens. Lett.* **2020**, *17*, 1488–1492. [CrossRef]

18. Xue, X.; Li, Z.; Ma, L.; Jia, Q.; Liu, R.; Fan, X. Investigating intrinsic degradation factors by multi-branch aggregation for real-world underwater image enhancement. *Pattern Recognit.* **2023**, *133*, 109041. [CrossRef]

19. Li, C.; Guo, C.; Ren, W.; Cong, R.; Hou, J.; Kwong, S.; Tao, D. An Underwater Image Enhancement Benchmark Dataset and Beyond. *IEEE Trans. Image Process.* **2020**, *29*, 4376–4389. [CrossRef] [PubMed]

20. Wu, J.; Liu, X.; Lu, Q.; Lin, Z.; Qin, N.; Shi, Q. FW-GAN: Underwater image enhancement using generative adversarial network with multi-scale fusion. *Signal Process. Image Commun.* **2022**, *109*, 116855. [CrossRef]

21. Wang, Y.; Guo, J.; Gao, H.; Yue, H. UIECˆ2-Net: CNN-based underwater image enhancement using two color space. *Signal Process. Image Commun.* **2021**, *96*, 116250. [CrossRef]

22. Chen, Y.; Ke, W.; Kou, L.; Li, Q.; Lu, D.; Bai, Y.; Wang, Z. Multiple Channel Adjustment based on Composite Backbone Network for Underwater Image Enhancement. *Int. J. Bio Inspired Comput.* 2023; *to be published*.

23. Chen, Y.; Li, H.; Yuan, Q.; Wang, Z.; Hu, C.; Ke, W. Underwater Image Enhancement based on Improved Water-Net. In Proceedings of the 2022 IEEE International Conference on Cyborg and Bionic Systems (CBS), Wuhan, China, 24–26 March 2023; pp. 450–454. [CrossRef]

24. Hambarde, P.; Murala, S.; Dhall, A. UW-GAN: Single-Image Depth Estimation and Image Enhancement for Underwater Images. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 5018412. [CrossRef]

25. Lin, P.; Wang, Y.; Wang, G.; Yan, X.; Jiang, G.; Fu, X. Conditional generative adversarial network with dual-branch progressive generator for underwater image enhancement. *Signal Process. Image Commun.* **2022**, *108*, 116805. [CrossRef]

26. Lu, J.; Li, N.; Zhang, S.; Yu, Z.; Zheng, H.; Zheng, B. Multi-scale adversarial network for underwater image restoration. *Opt. Laser Technol.* **2019**, *110*, 105–113. [CrossRef]

27. Chen, R.; Cai, Z.; Cao, W. MFFN: An Underwater Sensing Scene Image Enhancement Method Based on Multiscale Feature Fusion Network. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4205612. [CrossRef]

28. Liu, R.; Fan, X.; Zhu, M.; Hou, M.; Luo, Z. Real-World Underwater Enhancement: Challenges, Benchmarks, and Solutions Under Natural Light. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *30*, 4861–4875. [CrossRef]

29. Panetta, K.; Kezebou, L.; Oludare, V.; Agaian, S. Comprehensive Underwater Object Tracking Benchmark Dataset and Underwater Image Enhancement With GAN. *IEEE J. Ocean. Eng.* **2022**, *47*, 59–75. [CrossRef]

30. Li, S.; Song, W.; Qin, H.; Hao, A. Deep variance network: An iterative, improved CNN framework for unbalanced training datasets. *Pattern Recognit.* **2018**, *81*, 294–308. [CrossRef]

31. Li, G.; Zhang, M.; Li, J.; Lv, F.; Tong, G. Efficient densely connected convolutional neural networks. *Pattern Recognit.* **2021**, *109*, 107610. [CrossRef]

32. Gao, S.-H.; Cheng, M.-M.; Zhao, K.; Zhang, X.-Y.; Yang, M.-H.; Torr, P.H. Res2Net: A New Multi-Scale Backbone Architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 652–662. [CrossRef]

33. Liang, M.; Hu, X. Recurrent convolutional neural network for object recognition. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015. [CrossRef]

34. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShufflfleNet: An extremely effificient convolutional neural network for mobile devices. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6848–6856. [CrossRef]

35. Li, Z.; Peng, C.; Yu, G.; Zhang, X.; Deng, Y.; Sun, J. DetNet: Design Backbone for Object Detection. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 339–354. [CrossRef]

36. Fang, L.; Jiang, Y.; Yan, Y.; Yue, J.; Deng, Y. Hyperspectral Image Instance Segmentation Using Spectral–Spatial Feature Pyramid Network. *IEEE Trans. Geosci. Remote. Sens.* **2023**, *61*, 5502613. [CrossRef]

37. Zhang, J.; Wang, Y.; Wang, H.; Wu, J.; Li, Y. CNN Cloud Detection Algorithm Based on Channel and Spatial Attention and Probabilistic Upsampling for Remote Sensing Image. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5404613. [CrossRef]

38. Yi, S.; Liu, X.; Li, J.; Chen, L. UAVformer: A Composite Transformer Network for Urban Scene Segmentation of UAV Images. *Pattern Recognit.* **2023**, *133*, 109019. [CrossRef]

39. Zhao, Z.; Liu, Y.; Sun, X.; Liu, J.; Yang, X.; Zhou, C. Composited FishNet: Fish Detection and Species Recognition from Low-Quality Underwater Videos. *IEEE Trans. Image Process.* **2021**, *30*, 4719–4734. [CrossRef] [PubMed]

40. Yang, M.; Sowmya, A. An Underwater Color Image Quality Evaluation Metric. *IEEE Trans. Image Process.* **2015**, *24*, 6062–6071. [CrossRef] [PubMed]

41. Panetta, K.; Gao, C.; Agaian, S. Human-Visual-System-Inspired Underwater Image Quality Measures. *IEEE J. Ocean. Eng.* **2016**, *41*, 541–551. [CrossRef]