

# 光电工程

## Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊  
Scopus CSCD

### 基于残差和注意力网络的声呐图像去噪方法

赵冬冬, 叶逸飞, 陈朋, 梁荣华, 蔡天诚, 郭新新

#### 引用本文:

赵冬冬, 叶逸飞, 陈朋, 等. 基于残差和注意力网络的声呐图像去噪方法[J]. *光电工程*, 2023, 50(6): 230017.

Zhao D D, Ye Y F, Chen P, et al. Sonar image denoising method based on residual and attention network[J]. *Opto-Electron Eng*, 2023, 50(6): 230017.

<https://doi.org/10.12086/oe.2023.230017>

收稿日期: 2023-01-20; 修改日期: 2023-04-05; 录用日期: 2023-04-11

### 相关论文

#### 基于自适应BM3D的侧扫声纳图像散斑降噪

陈朋, 蔡烜伟, 赵冬冬, 梁荣华, 郭新新

*光电工程* 2020, 47(7): 190580 doi: 10.12086/oe.2020.190580

#### 分数傅里叶变换域的调频信号稀疏性研究

王硕, 郭勇, 杨立东

*光电工程* 2020, 47(11): 190660 doi: 10.12086/oe.2020.190660

更多相关论文见光电期刊集群网站 



<http://cn.ojournal.org/oe>



 OE\_Journal



Website

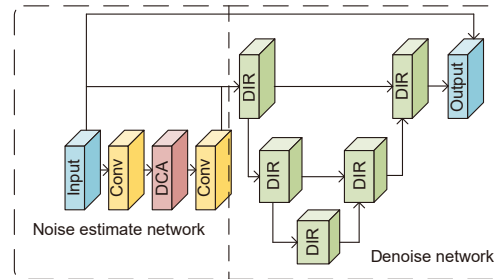
DOI: 10.12086/oe.2023.230017

## 基于残差和注意力网络的声呐图像去噪方法

赵冬冬<sup>1</sup>, 叶逸飞<sup>1</sup>, 陈朋<sup>1\*</sup>,  
梁荣华<sup>1</sup>, 蔡天诚<sup>1</sup>, 郭新新<sup>2</sup>

<sup>1</sup>浙江工业大学计算机科学与技术学院, 浙江 杭州 310023;

<sup>2</sup>中国科学院深海科学与工程研究所, 海南 三亚 572000



**摘要:** 前视声呐作为一种水下主动声呐设备常用于采集水下图像数据, 然而会受到水下噪声的影响导致图像质量下降。针对这一问题, 本文提出了一种基于密集残差和双通道注意力机制网络的前视声呐图像去噪方法。首先采用双通道注意力机制对声呐图像的通道信息进行提取, 统计声呐图像的全局信息, 输出声呐图像的噪声图; 密集残差块根据噪声图和声呐图像, 充分学习不同尺度上的特征信息, 经过多次学习和信息传递后输出干净声呐图像。针对前视声呐图像及其噪声特点, 模拟了前视声呐图像并添加瑞利分布的乘性噪声和高斯分布的加性噪声, 生成模拟数据集用于网络训练和性能评估。在模拟数据集和真实数据集的实验中表明, 本文方法能够有效去除噪声, 保留图像细节。

**关键词:** 前视声呐; 噪声模拟; 图像去噪; 通道注意力; 密集残差

中图分类号: U666.7

文献标志码: A

赵冬冬, 叶逸飞, 陈朋, 等. 基于残差和注意力网络的声呐图像去噪方法 [J]. 光电工程, 2023, 50(6): 230017

Zhao D D, Ye Y F, Chen P, et al. Sonar image denoising method based on residual and attention network[J]. *Opto-Electron Eng*, 2023, 50(6): 230017

## Sonar image denoising method based on residual and attention network

Zhao Dongdong<sup>1</sup>, Ye Yifei<sup>1</sup>, Chen Peng<sup>1\*</sup>, Liang Ronghua<sup>1</sup>, Cai Tiancheng<sup>1</sup>, Guo Xinxin<sup>2</sup>

<sup>1</sup>College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, Zhejiang 330063, China;

<sup>2</sup>Institute of Deep-sea Science and Engineering, Chinese Academy of Sciences, Sanya, Hainan 572000, China

**Abstract:** As a kind of underwater active sonar equipment, forward-looking sonar is often used to collect underwater image data. However, it will be affected by underwater noise, which leads to the degradation of image quality. Due to this situation, a forward-looking sonar image denoising method is proposed based on dense residuals and a dual-channel attention mechanism network. Firstly, the two-channel attention mechanism is adopted to extract the channel information of the sonar image, collect the global information of the sonar image, and output the noise map of the sonar image. Based on the noise image and sonar image, the dense residual block fully learns the feature information at different scales and outputs a clean sonar image after multiple learning and information transfer. Because of the forward-looking sonar image and its noise characteristics, the forward-looking

收稿日期: 2023-01-20; 修回日期: 2023-04-05; 录用日期: 2023-04-11

基金项目: 国家自然科学基金青年科学基金项目 (62001418); 浙江省自然科学基金项目 (LQ21F010011); 中国科学院战略性先导科技专项项目 (A类)(XDA22030302); 浙江省属高校基本科研业务费专项资金项目 (RF-C2019001)

\*通信作者: 陈朋, chenpeng@zjut.edu.cn。

版权所有©2023 中国科学院光电技术研究所

sonar image is simulated and the multiplicative noise of Rayleigh distribution and the additive noise of Gaussian distribution are added to generate a simulation dataset for network training and performance evaluation. Experimental results on the simulated data set and real data set show that the proposed method can effectively remove the noise and retain image details.

**Keywords:** forward looking sonar; image denoising; noise simulate; channel attention; dense residual

## 1 引言

随着人类社会经济的发展, 在海洋区域上人们也在积极探索, 海洋中所蕴含的丰富的能源、矿产和生物资源具有极高的价值, 因此如何获取水下信息数据对于海洋科学和海洋经济的发展至关重要。前视声呐作为一种用于探测水下目标并测定其速度、位置、运动方向的设备, 在海底勘探、海洋渔业资源捕获、军事目标检测等方面都有广泛应用。

由于电磁波在水中传播时大部分能量被海水所吸收, 传播距离有限, 传统的光学成像系统难以获得远距离的清晰图像, 前视声呐虽然有足够的探测距离。然而在水下声呐图像受到的噪声干扰严重, 分辨率较低且容易出现信息缺失<sup>[1-2]</sup>, 因此在对前视声呐图像进行去噪时, 需要在去除噪声的同时, 尽可能地保留声呐图像的纹理、边缘等细节信息。

传统去噪算法的类型大致称为: 空间域滤波、变换域阈值滤波、统计模型、稀疏表示方法、字典学习方法和混合方法等<sup>[3]</sup>。其中空间域滤波包括局部滤波和非局部滤波, 常用的降噪局部滤波器有高斯滤波器、最小均方滤波器、双边滤波器、维纳滤波器、苏珊滤波器等<sup>[4]</sup>。局部滤波器的实现思想是基于像素之间的相关性实现的, 在高噪声水平的图像上表现较差。为了解决这个问题, Buades<sup>[5]</sup>等人引入了一种非局部均值滤波器, 该方法以非局部的方式在整个图像区域中利用像素之间的相似性来去噪。Wu<sup>[6]</sup>等人提出了一种基于稀疏表示的声呐图像去噪方法, 对原图进行DCT字典分解重构和OMP去噪。Jiang<sup>[7]</sup>等人提出了一种利用分析模型进行稀疏分解的去噪方法, 该方法对于侧扫声呐图像的散斑噪声具有良好的去噪效果。Chen<sup>[8]</sup>等人提出了一种基于固定水域噪声模型的水声图像去噪方法, 在对声呐图像去噪的同时保持一定的图像边缘信息和纹理信息。Stoljescu-Crisan<sup>[9]</sup>等人提出了一种在小波域上的声呐图像处理算法, 填补了去噪时缺失像素产生的缺口。范等人<sup>[10]</sup>和陈等人<sup>[11]</sup>都提出了一种基于自适应三维块匹配滤波(BM3D)的侧

扫声呐图像散斑降噪方法, 有效地降低侧扫声呐图像中的散斑噪声获得了良好的视觉效果。然而传统去噪方法往往是在熟悉了噪声的规律下才能产生较好的去噪效果, 对于一些复杂的或者强度水平不同噪声效果往往并不理想。

随着计算机硬件设备性能的飞速提升, 基于深度学习的去噪方法也在蓬勃地发展, 卷积神经网络在光学领域上的去噪具有非常大的发展前景。Zhang<sup>[12]</sup>等人为了能够使用单一模型去除不同噪声水平的图像, 提出了基于下采样子图像的网络FFDNet, 该网络在推理速度和去噪性能之间的平衡性比较好。DnCNN<sup>[13]</sup>和ASCN<sup>[14]</sup>提供了一种新的思路, 通过学习真实图像和训练图像的残差分布, 对降噪问题进行建模。ResNet<sup>[15]</sup>提出了残差思想, 解决了深度CNN训练困难的问题, 许多人将其应用创新在图像处理领域<sup>[16-17]</sup>, Ledig<sup>[18]</sup>等人提出了基于对抗生成网络(GAN)和残差网络(ResNet)的方法用于图像的去噪和超分辨率, 在光学图像上能够恢复较真实的图像纹理。王晓娜<sup>[19]</sup>等人提出了一种分通道反卷积的方法, 该方法首先对原始图像进行颜色空间转换, 然后对不同颜色通道进行分别处理, 相比传统的反卷积, 该方法在保证清晰度的同时, 降低了图像噪声, 并且缩短了运行时间。Cheng<sup>[20]</sup>等人通过学习特征空间中的一组重构基来训练一个能够分离信号和噪声的UNet<sup>[21]</sup>结构网络, 选择相应的信号子空间基并将输入投影到该子空间中, 利用全局结构信息来保持低亮度和弱纹理的区域结构。随着注意力(Attention)机制在自然语言处理领域的成功应用, 也有人将其应用到了图像处理中来实现不同的任务<sup>[22-24]</sup>。在图像去噪方面, Wang<sup>[25]</sup>等人将空间注意力机制和通道注意力机制融合在一起, 提出了一种通道和空间注意力神经网络(CS-ANN)用于图像去噪, 采用了对称跳跃连接和残差块避免梯度消失和浅层特征的损失。Tian<sup>[26]</sup>等人在注意力引导机制的基础上提出了一种图像去噪卷积神经网络, 网络通过稀疏块、特征增强块、注意力机制块和图像重构块串联构成, 提取复杂背景中隐藏的噪声并去除。CBDNet<sup>[27]</sup>



由噪声估计子网络和噪声去除子网络两个子网络构成。在噪声估计阶段, 使用了连续的多个卷积层对输入的原图像进行噪声提取; 在噪声去除阶段, 将提取到的噪声和原图一起输入到 Unet<sup>[21]</sup> 中, 输出去噪图像。目前基于神经网络的图像去噪方法大多都是基于光学图像上的去噪, 声学图像上的应用研究并不像光学如此丰富。并且现有方法在网络加深时收敛速度容易下降, 特征图经过层层传递容易丢失部分信息, 导致全局信息利用不完全, 且特征提取方式单一, 导致训练效果变差。

由于声呐图像的特殊性, 公开的声呐图像数据集十分匮乏且获取困难, 如何将深度学习的方法应用到声呐图像的去噪上是一个不小的挑战。Jensen<sup>[28]</sup> 等人提出了一种利用声学原理设计的一个超声系统 FieldII, 该系统可以模拟超声成像系统, 通过配置不同类型的超声传感器, 计算发射声场和回波声场来生成图像, 可以应用在医疗图像上, 前视声呐也是一种声学设备, 因此可以使用 FieldII 来模拟生成前视声呐数据。

本文提出了一种基于密集残差和双通道注意力机制网络 (dense in residual and dual channel attention network, DIRANet) 用于前视声呐图像去噪。该算法首先使用双通道注意力 (dual channel attention, DCA) 模块提取声呐图像的通道信息, 捕获全局信息并输出声呐图像的噪声图, 之后采用密集残差 (dense in residual, DIR) 模块, 对原图和噪声图均采用编码-解码结构进行图像融合重建, 充分学习不同尺度上的特征信息, 经过多次学习和信息传递之后, 生成干净

图像。最后使用 FieldII 生成的模拟前视声呐数据集并添加模拟噪声的数据集以及真实声呐数据集进行去噪对比实验, 结果表明本文算法在各项指标上均优于对比的图像去噪算法。

## 2 DIRANet

本文提出了密集残差和双通道注意力机制网络 (DIRANet) 用于前视声呐的图像去噪, 其中设计了双通道注意力机制 (DCA) 模块用于噪声估计以及密集多尺度残差块 DIR 用于噪声去除, 网络整体由噪声估计和噪声去除两个阶段构成。在噪声估计阶段利用 DCA 模块提取图像通道之间的联系, 统计全局信息, 获得更多和噪声有关的信息, 并输出噪声估计图; 噪声去除阶段利用 DIR 模块对噪声估计图和原图像进行不同尺度上的特征高效提取, 提升网络的接受域范围, 获取更多的特征信息, 最终输出去噪后的前视声呐图像。

图 1 为本文所提出的 DIRANet 网络结构图, 该网络主要分为两个子网络。第一阶段为噪声估计子网络, 第一个块为  $3 \times 3$  的卷积层和 ReLU 激活函数构成, 接着是一个 DCA 模块, 用于增强特征图的通道联系, 统计全局信息, 最后经过一个卷积层和 ReLU 激活函数, 输出噪声估计水平图。第二阶段为噪声去除子网络, 该网络主体结构为编码器-解码器结构, 编码器和解码器由 DIR 模块构成, DIR 模块通过块内构建的分层残差结构, 对不同层的特征图进行多尺度的特征提取, 提升网络的接受域范围, 最终输出噪声去

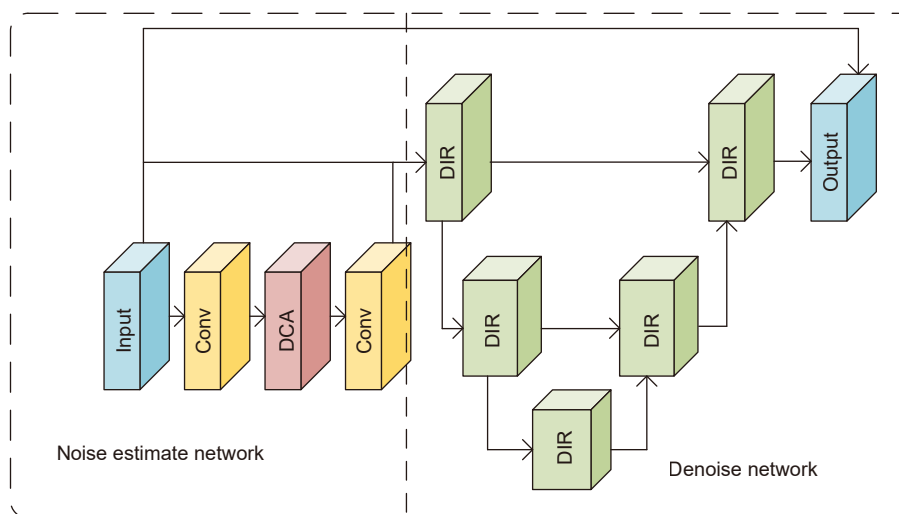


图 1 DIRANet 整体结构图

Fig. 1 DIRANet overall structure diagram

除后的图像。

### 2.1 噪声估计子网络

在神经科学中视觉皮层神经元的感受野大小是由刺激调节的, 基于这种思想, 本文提出了一种基于分离、融合过程的双通道注意力机制 (DCA) 模块用于构建噪声估计子网络, DCA 模块结构如图 2 所示。在分离阶段, 使用不同卷积核大小来自适应地接受不同规模的信息, 并分别将这些信息分流到下一层神经元中, 这些不同的分支融合之后经过全局平均池化操作后融合生成一个全局的注意力权重信息, 之后用一维卷积生成一维向量, 相比于全连接层, 一维卷积层在处理通道注意力权重时, 可以直接对通道维度上的特征进行卷积操作, 提取特征之间的相互依赖关系, 而不需要将整个通道特征进行展平并传入全连接层中, 获取更多信息的同时减少了计算量。

DCA 将输入的特征图  $U \in R^{C \times H \times W}$  分离成两组通道数相同的特征图  $X_1, X_2 \in R^{C \times H \times W}$ , 对其分别进行  $3 \times 3$  和  $5 \times 5$  卷积核大小的卷积得到  $X'_1 \in R^{C/2 \times H \times W}$  和  $X'_2 \in R^{C/2 \times H \times W}$ , 获得不同感受野的特征图。

将  $X'_1$  和  $X'_2$  进行连接操作得到新的特征图  $F' \in R^{C \times H \times W}$ , 随后, 我们对特征图  $F'$  进行分组提取权重。首先通过二维全局平均池化 (AverageGlobalPooling) 得到  $F'_1 \in R^{C \times 1 \times 1}$ , 将  $F'_1$  压缩 (squeeze) 到二维  $F'_2 \in R^{C \times 1}$ ,

然后将其进行矩阵转置 (transpose) 得到  $F'_{2t} \in R^{1 \times C}$ , 对  $F'_{2t}$  进行一维卷积操作, 将结果进行维度拓展 (unsqueeze) 与矩阵转置得到总体权重矩阵  $w \in R^{C \times 1 \times 1}$ 。

$$w = \text{unsqueeze}(\text{conv}(\text{transpose}(\text{squeeze}(\text{avgpool}(F'_1))))). \quad (1)$$

对  $w$  进行两次相同的操作分别得到特征图  $X'_1$  对应的通道权重矩阵  $w_1 \in R^{C/2 \times 1 \times 1}$  和特征图  $X'_2$  对应的通道权重矩阵  $w_2 \in R^{C/2 \times 1 \times 1}$ 。

将特征图  $X'_1$ 、 $X'_2$  分别和权重  $w_1$ 、 $w_2$  相乘, 并将  $X'_1$ 、 $X'_2$  进行连接, 得到最终输出的特征图  $U_{out}$ 。整体流程表示如式 (2):

$$U_{out} = \text{concat}\left(\sum w_1 \cdot X'_1, \sum w_2 \cdot X'_2\right). \quad (2)$$

DCA 结构有如下优势: 1) 相比一般的通道注意力机制的模块, 本结构设计了一种双路的通道注意力机制, 主要通过使用两个不同尺寸的卷积核实现, 通过  $3 \times 3$  和  $5 \times 5$  的两个不同大小的卷积核构成两条路径, 这两条路径能够获取不同尺寸的信息, 提升网络的信息提取能力。2) 使用一维卷积而不是传统的全连接层生成通道特征图, 因为一维卷积的参数数量比全连接层要少得多, 并且可以直接在通道维度上进行卷积操作, 不需要将整个通道特征作为参数传入。并且一维卷积可以自适应地学习相邻通道之间的权重, 增强通道之间的联系。

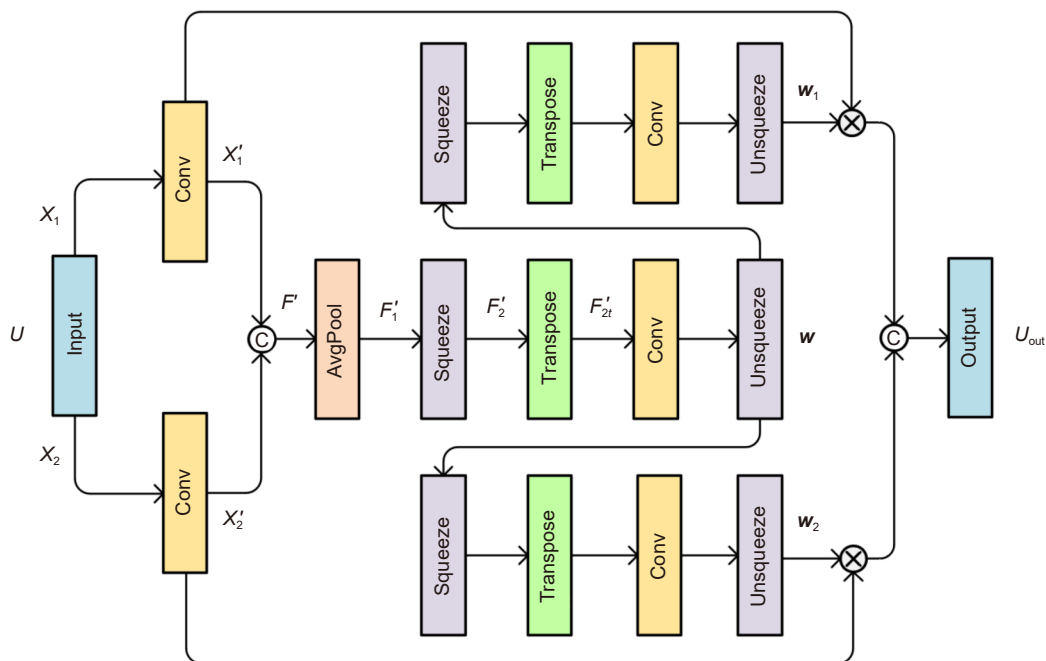


图 2 DCA (dual channel attention) 模块

Fig. 2 DCA (dual channel attention) module

### 2.2 噪声去除子网络

噪声去除子网络将上层网络输出的噪声估计图以及原始图像作为输入, 进行噪声去除。网络整体结构由编码器和解码器堆叠构成, 每个模块分层提取图像的浅层和深层信息, 并将不同深度的信息进行合并处理, 最终输出去噪后的图像。

本文提出 dense in residual (DIR) 结构如图 3 所示。为了能够在更细粒度的水平上提高多尺度的表示能力, 首先将输入的特征图依据通道数划分成  $s$  个特征图子集, 和原特征图相比, 每个新子集都有相同的特征图大小, 但是通道数为原特征图通道数的  $1/s$ , 每一个  $x_i$  都对应一个卷积层, 为了匹配这些特征图子集, 将卷积层中具有  $n$  个通道的  $3 \times 3$  的卷积核替换成了  $s$  个通道数为  $m$  的  $3 \times 3$  卷积核组, 其中  $n = s * m$ , 其中每一个卷积核  $\text{Conv}_j$  都可以都从特征图子集  $x_i$ ,  $\{x_i; i \leq j\}$  中接受特征信息, 并且每一个分割的  $x_i$  进行一次卷积操作之后, 都可以输出一个比  $x_j$  更大的接受域, 除了第一组卷积核只接受其对应的输入特征图, 之后的每一组卷积核都接受之前所有卷积核的输出和其对应的输入特征图, 直至处理完所有的输入特征图。每个卷积层之间的紧密连接, 使得每一层学习到的特征能够较快地传播给下一层, 其学习曲线会比较平滑, 经过这样层层的不同大小的特征组合提取, 可以将信息进行充分地融合。

DIR 结构使用了残差结构中密集连接的设计, 具体而言, 将  $1 \times 1$  卷积层的输出拆分成 4 个小组,  $3 \times 3$  卷积层替换成了 3 个较小的  $3 \times 3$  卷积对特征图进行分组卷积, 每一个小卷积的结果又会输入到其后面的每一个小卷积上, 将所有结果进行连接操作, 加上恒等映射层, 输出最终结果:

$$y_i = \begin{cases} x_i & i = 1 \\ \text{Conv}_i(x_i) & i = 2 \\ \text{Conv}_i(y_1, y_2, \dots, y_{i-1}, x_i) & 2 < i \leq g \end{cases}, \quad (3)$$

其中,  $y_i$  表示每一小组经过卷积后输出的特征图,  $\text{Conv}_i$  表示卷积操作, 其中每一个卷积层所接受的特征图大小都不相同, 融合了多种尺度的特征图。

DIR 模块有如下优势: 1) 相比一般的残差模块, 本文将原有的  $3 \times 3$  卷积层切割成了若干组卷积核数量相等的  $3 \times 3$  卷积, 分组卷积可以减少参数量提升训练速度, 并且前一组卷积都会将其结果作为输入传递到后一组卷积中进行处理, 这种方式提升了之后每一组卷积的接收信息量, 增加了信息处理的能力。2) 在更细粒度上设计了残差和密集的连接方式, 每一层之间的紧密连接使得信息的传递更加的快速和流畅, 每一组卷积可以接收前几层所学到的知识, 提升网络的特征提取能力。

## 3 数据集

### 3.1 模拟前视声呐数据集

本文的模拟前视声呐数据使用 FieldII 来模拟生成。设置中心频率为 350 kHz, 采样频率为 1.4 MHz, 圆弧形发射阵, 129 个换能器, 520 条波束。为了让模拟的前视声呐图像更加贴近真实场景, 模拟了不同形状的物体并将其放置在声呐场中的不同位置, 波束形成之后得到原始声呐图像, 将极坐标系下的图像转化到笛卡尔坐标系, 最终得到了扇形的前视声呐图像如图 4 所示。

声呐图像的噪声主要为乘性散斑噪声<sup>[29-31]</sup>, Yu 等人<sup>[29]</sup>指出散斑噪声是一种乘性的局部相关的噪声, 在医学超声图像中都存在, 对于图像的质量会造成不

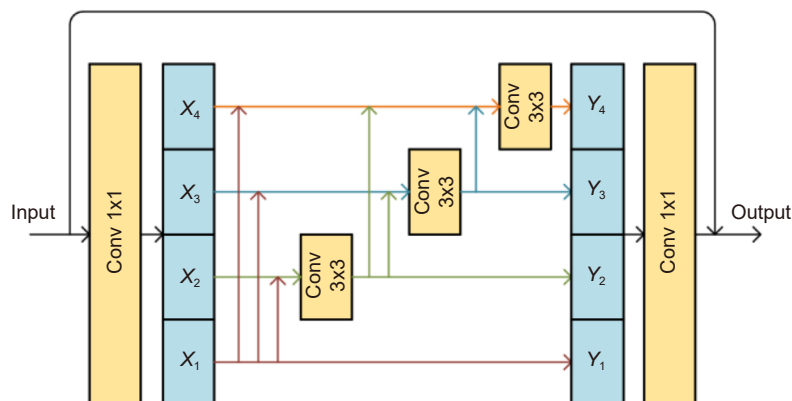


图 3 DIR (dense in residual) 结构

Fig. 3 DIR (dense in residual) structure

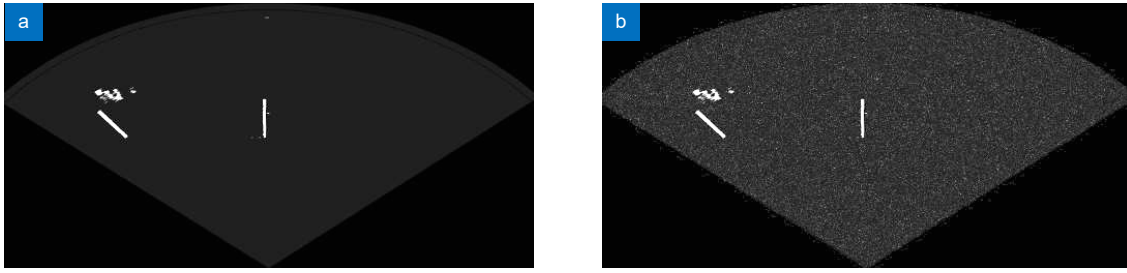


图4 模拟前视声呐图像。(a) 原始图像; (b) 噪声图像

Fig. 4 Simulated forward-looking sonar images. (a) Original image; (b) Noisy image

小的影响除了散斑噪声之外, Karabchevsky 等人<sup>[30]</sup>也指出了散斑噪声是存在于激光成像、雷达、SAR 和超声波成像等相干成像中的一种颗粒状噪声。散斑噪声是一种乘法噪声, 即它与该区域的局部强度成正比, 基于这种特性, 可以用具有指数分布的乘法模型来描述这种噪声  $R$ 。

$$R = A_x + A_y, \quad (4)$$

其中,  $R$  表示瑞利分布的噪声,  $A_x$  和  $A_y$  表示一对具有高斯分布、均值为零、方差相同的绝对独立的随机变量。当一个随机二维向量的两个分量呈独立的、均值为零, 有着相同的方差的正态分布时, 这个向量模呈瑞利分布。因此这两个高斯分布叠加之后的分布也就是瑞利分布。因此使用瑞利分布的模型可以模拟乘性散斑噪声。

水下的环境噪声对图像质量也有一定的影响, 但是相比散斑噪声对图像的影响要小得多而环境噪声可以描述为加性的高斯噪声<sup>[32-33]</sup>, 在这里本文将前视声呐带噪声的图像模型用如式 (5) 表示:

$$I_R(x, y) = I(x, y) \cdot R(\sigma(t)) + N_G(x, y), \quad (5)$$

其中,  $I_R$  表示含噪声图,  $I$  表示原声呐图像,  $R(\sigma(t))$  表示均值为 0、方差为  $\sigma(t)$  的瑞利分布噪声, 其方差随时间在区间  $[0.1, 1.5]$  内均匀分布, 可用来模拟非平稳环境下的噪声,  $N_G$  表示均值为 0 方差为 1 的加性高斯噪声。总体上模拟图像添加了乘性的瑞利噪声和加性的高斯噪声, 噪声的方差会随着时间的改变而改变, 这样的不稳定噪声图像可以使网络学习到更多的噪声特征并增强网络对不同噪声环境的适应能力和鲁棒性。

本文使用的模拟前视声呐图像如图 4(a) 所示, 图 4(b) 为添加噪声后的图像, 最终生成了 10000 份模拟前视声呐图像噪声对, 将其中的 7000 对作为网络的训练集, 剩下的取 2000 对作为验证集, 1000 对作为测试集。

### 3.2 真实声呐数据集

本文选择 Giacomo<sup>[34]</sup> 等人通过前视声呐采集到的数据作为真实数据集, 该数据集使用型号为 LBV300-5 的探测车辆, 车辆上配有差分全球位置系统 (DGPS)、指南针以及 Teledyne BlueView 的多波束前视声呐设备。车在  $30^{\circ}01'30.1''S$   $52^{\circ}06'24.0''W$  位置, 维持在水下 0.3 米的深度和水上 3 米的高度行驶, 前视声呐设备以采集距离 50 m, 开角 130 度的形态收集数据, 总共收集了 2894 幅声学图像, 图像里包含石头、码头、河岸、船体等物体的声学信息。

## 4 实验结果与分析

### 4.1 全参考图像质量评估

对于模拟生成的前视声呐图像噪声对, 采用峰值信噪比 (peak signal to noise ratio, PSNR) 来评价噪声去除效果, 使用结构相似性 (structural similarity, SSIM) 来评价去噪图像的信息完整性。

#### 4.1.1 峰值信噪比

给定一张大小为  $m \times n$  的原图像  $I$  与经过处理的图像  $K$ , PSNR (公式中简称为  $P_{\text{PSNR}}$ ) 定义为:

$$P_{\text{PSNR}} = 10 \cdot \log_{10} \left( \frac{2^B - 1}{M_{\text{MSE}}} \right) \quad (6)$$

其中,  $B$  为图像像素点表示的位数若图像为 8 位, 则其所能表示的最大像素值为  $2^8 - 1 = 255$ , 均方差定义为 MSE (公式中简称为  $M_{\text{MSE}}$ ):

$$M_{\text{MSE}} = \frac{1}{m \times n} \sum_{i=1}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2. \quad (7)$$

PSNR 的值越大, 表示去噪后的图和原图越接近, 也就是去噪效果越好。

#### 4.1.2 结构相似性

给定两个图像  $x$  和  $y$ , 两张图像的结构相似性



(SSIM, 公式中简称为  $S_{SSIM}$ ) 定义如式 (8) 所示:

$$S_{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (8)$$

其中,  $\mu_x$  是图像  $x$  的平均值,  $\mu_y$  是图像  $y$  的平均值,  $\sigma_x^2$  是图像  $x$  的方差,  $\sigma_y^2$  是图像  $y$  的方差,  $\sigma_{xy}$  为  $x$  和  $y$  的协方差。  $c_1 = (k_1L)^2, c_2 = (k_2L)^2$ , 这两个值是用来维持数值稳定的常数。  $L$  是图像像素值的动态范围。  $k_1 = 0.01, k_2 = 0.03$ 。 SSIM 值越大则表示两张图像越相似。

#### 4.2 无参考图像质量评估

对于没有真实干净图像的声呐图像, 采用一种无参考的空间域图像质量评估算法 *Brisque*<sup>[35]</sup> 来评价去噪效果。该算法首先提取自然场景统计信息, 计算局部归一化亮度, 然后将该值拟合到广义高斯分布 (GGD) 来计算特征向量, 最后放到支持向量机中进行训练, 生成图像质量得分。

#### 4.3 模拟声呐图像对比实验

BM3D<sup>[10]</sup> 算法首先对每个参考块进行相似块匹配 (Block-Matching) 并得到一个三维组合, 然后对该组合进行三维协同变换滤波 (3D-Transform), 最后对滤波结果进行整合, 生成最终去噪结果图。

DnCNN 网络<sup>[13]</sup> 在 VGG 网络结构的基础上进行修改, 整体上是卷积层 (Conv)、批量归一化层 (BN) 和激活函数层 (ReLU) 级联并不断重复的结构, 模型内部不使用跳跃连接, 但是网络的输出是残差图像, 即噪声图像减去干净图像后的结果, 因此将噪声图像减去残差图像即可得到干净图像。 CBDNet<sup>[27]</sup> 由噪声估计-噪声去除两阶段构成, 对于不同的噪声, 通过两个子网络强化对噪声的提取和去除, 提高去噪效果。 CNCL<sup>[36]</sup> 是一种基于深度学习中的生成对抗网络, 该网络中的生成器生成干净图像, 判别器通过互补的方式学习图像数据集中的图像与噪声两种信息进行更加精确的判别。 DeamNet<sup>[37]</sup> 是一种基于模型的去噪方法, 利用自适应一致先验性模型的推理过程指导神经网络的优化, 最终建立一个端到端去噪网络, 将模型知识和深度学习相结合。

本文分别采用 BM3D<sup>[10]</sup>, DnCNN<sup>[13]</sup>, CBDNet<sup>[27]</sup>, CNCL<sup>[36]</sup>, DeamNet<sup>[37]</sup> 和本文提出的算法对模拟声呐图像噪声对进行去噪实验, 对测试集上的所有图片进行测试验证。实验结果如图 5 和表 1 所示, 图中左侧的红色方框表示的是需要进行细节放大的区域, 红色箭头指向的是该红色方框放大之后的细节显示区域。表 1 为 BM3D<sup>[10]</sup>, DnCNN<sup>[13]</sup>, CBDNet<sup>[27]</sup>, CNCL<sup>[36]</sup>,

DeamNet<sup>[37]</sup> 和本文方法处理模拟前视声呐图像噪声对得到的指标对比, 其中标注粗体的数值表示其实验结果最好的。首先对比传统方法和神经网络方法, 观察图 5 我们也可以发现, BM3D 在去噪时去除了过多的图像信息, 因为本文所用的声呐图像包含非线性噪声且图中有大量小块, 导致其块匹配过程会出现一些错误匹配的情况, 因此损失了不少的图像信息和细节对图像的破坏比较大, 去噪图像比较模糊; DnCNN、CNCL 和 CBDNet 对于复杂噪声和非线性噪声的处理能力较弱, 没有注意力机制的引导不能很好地关注到全局噪声, 因此不能很好地去除噪声; DeamNet 因为其自适应一致性先验的局限性, 会将一些图像信息错误判断成噪声信息, 因此在去除噪声时会出现原图有些模糊的情况; 综合来看本文提出的方法使用双通道注意力机制捕获了全局噪声信息, 并且通过密集残差块提升了图像的信息传递能力, 保留了较多的图像非噪声信息, 因此去噪效果最好。观察表 1, 五种神经网络方法相比传统算法的 BM3D 均有比较大的优势, PSNR 和 SSIM 值均比 BM3D 要高, 也就是去噪效果和图像细节保存程度均是神经网络算法表现较好。其次对比五种神经网络方法, 从直观视觉来看, 本文提出算法的去噪效果在几种方法里去噪效果最好。观察表 1 可以发现, 本文提出的算法具有比较大的优越性, 在对比实验中 PSNR 和 SSIM 值均为最佳结果。在 PSNR 上, 相比 DnCNN 有约 6.69% 的提升, 相比 CBDNet 有约 5.42% 的提升, 相比 CNCL 提升了约 5.79%, 相比 DeamNet 提升了约 2.73%; 在 SSIM 上, 相比 DnCNN 有约 3.45% 的提升, 相比 CBDNet 有约 7.62% 的提升, 相比 CNCL 提升了约 8.10%, 相比 DeamNet 提升了约 9.02%。

#### 4.4 真实声呐图像对比实验

本文将 BM3D<sup>[10]</sup>, DnCNN<sup>[13]</sup>, CBDNet<sup>[27]</sup>, CNCL<sup>[36]</sup>, DeamNet<sup>[37]</sup> 和本文提出的算法作为对比算法, 在所有的真实声呐数据上进行去噪实验。实验结果如图 6 和表 2 所示, 图中左侧的红色方框表示的是需要进行细节放大的区域, 红色箭头指向的是该红色方框放大之后的细节显示区域。表 2 为 BM3D<sup>[10]</sup>, DnCNN<sup>[13]</sup>, CBDNet<sup>[27]</sup>, CNCL<sup>[36]</sup>, DeamNet<sup>[37]</sup> 和本文算法处理真实声呐图像噪声对得到的图像质量评价指标对比, 其中标注粗体的数值表示其实验结果最好。观察图 6, BM3D 算法因为声呐图像中包含的非线性噪声和图中大量的小块信息, 干扰了其块匹配和滤波



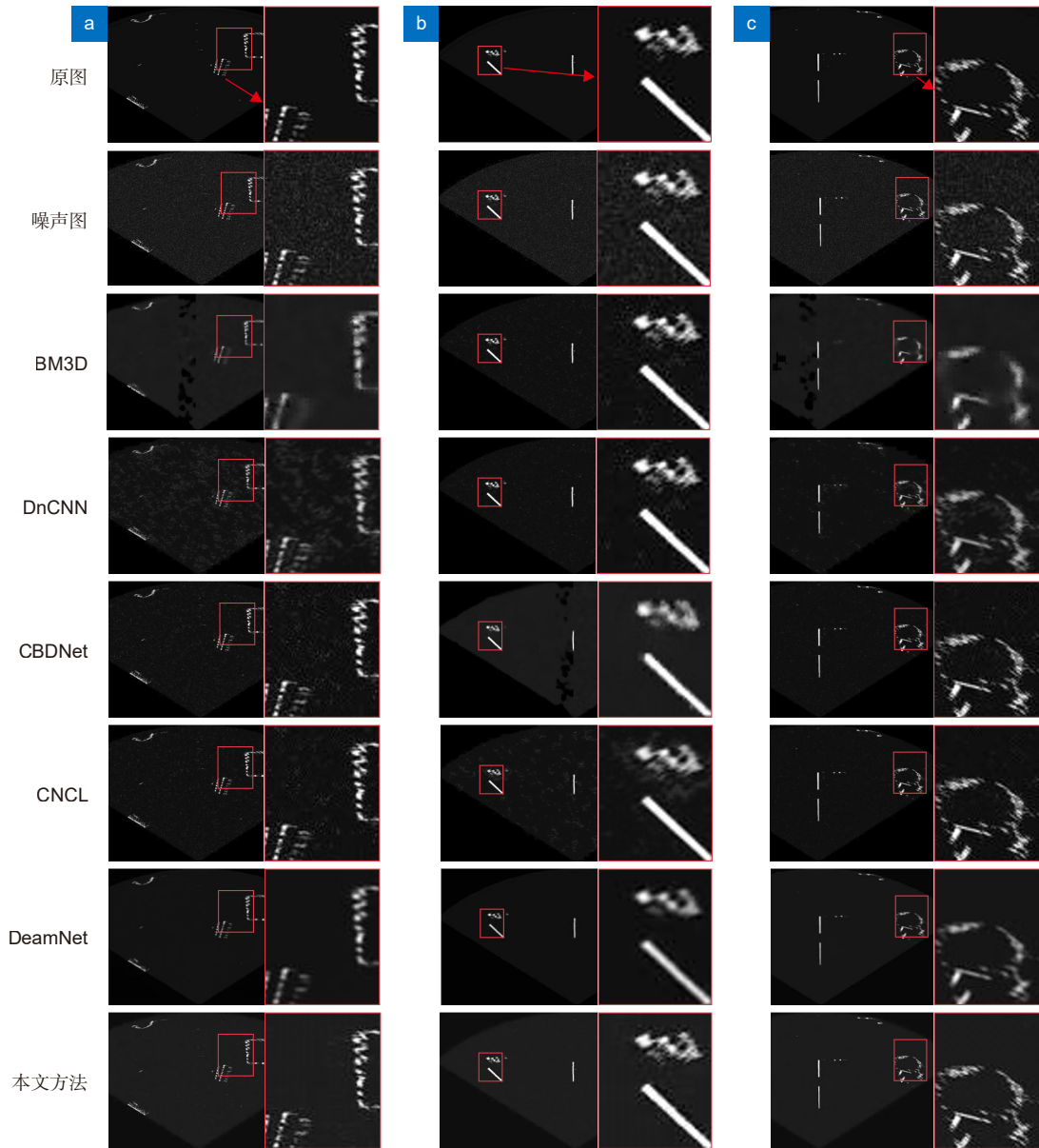


图 5 模拟前视声呐图像去噪结果, (a) (b) (c) 表示三种不同的图像

Fig. 5 Simulated forward-looking sonar image denoising results. (a) (b) (c) Represents three different images

表 1 模拟声呐图像去噪结果

Table 1 simulated sonar image denoising results

方法	PSNR	SSIM	Brisque
BM3D <sup>[10]</sup>	29.84	0.7693	104.18
DnCNN <sup>[13]</sup>	34.25	0.8168	95.27
CBDNet <sup>[27]</sup>	34.66	0.7852	98.82
CNCL <sup>[36]</sup>	34.54	0.7817	94.68
DeamNet <sup>[37]</sup>	35.57	0.7751	95.70
本文	<b>36.54</b>	<b>0.8450</b>	<b>91.13</b>

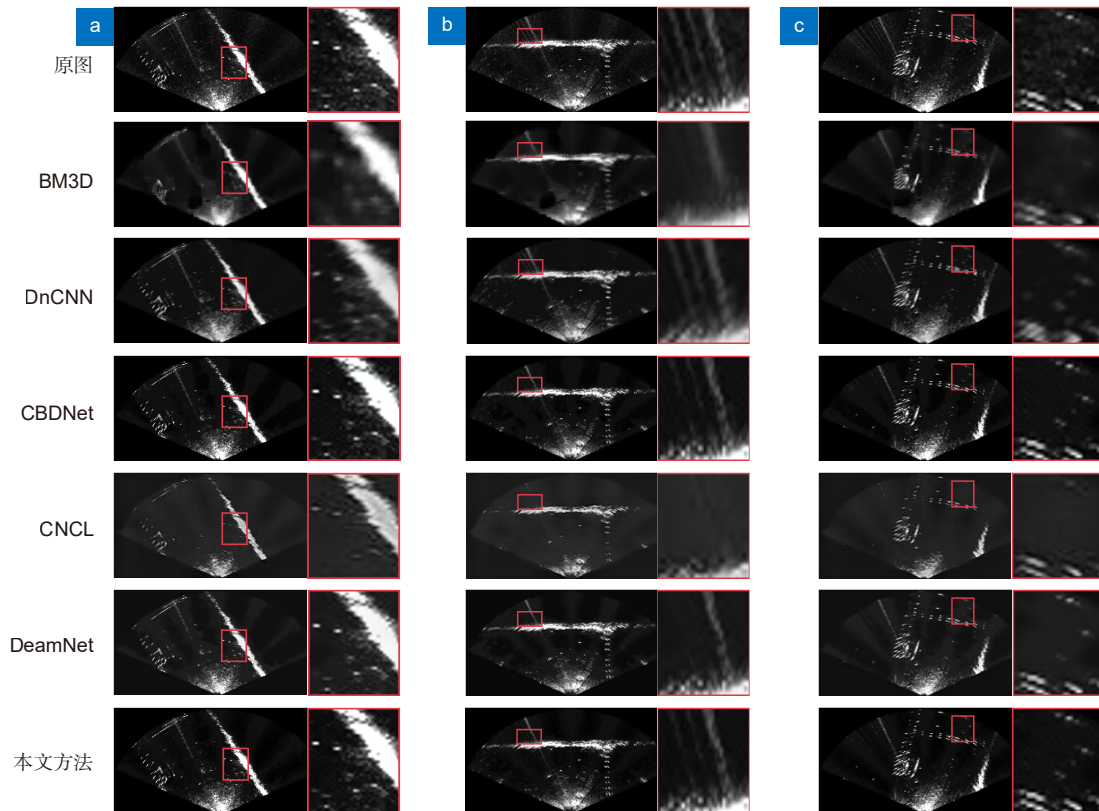


图6 真实前视声呐图像去噪结果, (a) (b) (c) 表示三种不同的图像

Fig. 6 Real forward-looking sonar image denoising results. (a) (b) (c) Represents three different images

过程, 导致有较多的图像信息和边缘细节被模糊化了, 去噪范围过大, 整体去噪效果较差; DnCNN、CBDNet、CNCL 和 DeamNet 缺乏对全局噪声信息的关注, 在去除噪声时对于图像保留欠佳, 因此对于图像信息的边界部分处理较为模糊甚至会损失部分图像细节, 图像明暗对比度相对于原图也被弱化了。总体而言本文算法使用双通道注意力机制捕获全局噪声信息的同时, 使用密集残差块提升网络的信息传递能力避免损失图像信息, 因此在去除噪声时能够较好地保留了细节信息, 总体去噪效果最好。观察表 2, 在指标上本文提出的算法 *Brisque* 指数最小, 表现最好。相比 DnCNN 降低了约 7.55%, 相比 CBDNet 降低了约 5.40%, 比 CNCL 降低了约 11.0%, 比 DeamNet 降低了约 14.7%。

#### 4.5 网络各模块有效性验证

##### 4.5.1 DCA 模块的有效性

为了证明 DCA 模块中的双通道注意力机制的有效性, 在本文方法整体结构不变的基础上, 改变 DCA 模块中的路径数量, 在模拟数据集和真实数据集上进行对比实验。最终的实验结果评价指标 PSNR、

SSIM 和 *Brisque* 数值如表 3 所示。对比表 3 中的数据可以看出, DCA 相比单路径通道注意力模块能够获得更加丰富的噪声信息, 使得网络模型的性能更好, 因此验证了本文设计的 DCA 模块的有效性。

##### 4.5.2 DIR 模块的有效性

为了证明 DIR 模块中的多尺度卷积机制的有效性, 在本文方法整体结构不变的情况下, 对比是否使用 DIR 块内分组卷积的情况, 在模拟数据集和真实数据集上进行对比试验。最终的实验结果评价指标 PSNR、SSIM 和 *Brisque* 数值如表 4 所示。对比表 4 中的数据可以看出, DIR 模块使用分组卷积和密集残差结构能够增强其细粒度上的多尺度特征提取能力, 避免信息的丢失, 提升网络的去噪能力, 因此验证了本文设计的 DIR 模块的有效性。

#### 4.6 消融实验

为了验证本文提出的 DIR 模块和 DCA 模块的有效性和通用性, 分别在 CBDNet<sup>[27]</sup> 和 CNCL<sup>[37]</sup> 网络中添加了本文提出的模块进行对比实验。实验在模拟数据集和真实数据集上进行, 最终的实验结果 PSNR、SSIM 和 *Brisque* 如表 5 所示, 标注粗体的数值表示

表 2 真实声呐图像去噪评价指标结果表

Table 2 Real sonar image denoising evaluation index results

方法	Brisque
BM3D <sup>[10]</sup>	53.81
DnCNN <sup>[13]</sup>	43.57
CBDNet <sup>[27]</sup>	42.58
CNCL <sup>[36]</sup>	45.27
DeamNet <sup>[37]</sup>	47.21
本文	<b>40.28</b>

表 3 不同注意力机制的比较

Table 3 Comparison of different attention mechanisms

模型	模拟数据集			真实数据集
	PSNR	SSIM	Brisque	Brisque
单路注意力	35.68	0.8276	94.19	41.61
DCA	36.54	0.8450	91.13	40.28

表 4 不同残差块的比较

Table 4 Comparison of different residual block

模型	模拟数据集			真实数据集
	PSNR	SSIM	Brisque	Brisque
普通残差块	35.98	0.8367	95.26	41.34
DIR	36.54	0.8450	91.13	40.28

表 5 消融实验结果对比

Table 5 Comparison of ablation results

方法	模拟数据集			真实数据集
	PSNR	SSIM	Brisque	Brisque
CBDNet	34.66	0.7852	98.82	42.58
CBDNet+DCA	35.87	0.7961	97.48	41.89
CBDNet+DIR	35.43	0.8257	98.31	42.04
CNCL	34.54	0.7817	94.68	45.27
CNCL+DCA	34.84	0.7746	94.42	44.35
CNCL+DIR	35.02	0.7993	92.88	43.79
本文	<b>36.54</b>	<b>0.8450</b>	<b>91.13</b>	<b>40.28</b>

效果最好。

对比表 5 的数据可以看出, DCA 模块可以通过多路注意力机制捕获不同尺度的信息, 获取更多的全局信息, 从而更好地提取噪声信息, 因此验证了本文提出的双通道注意力机制模块的有效性和通用性; 通过表 5 的数据可以得知 DIR 模块的多尺度卷积机制提升了整体模块感受野的信息收益, 并且利用细粒度上的密集连接机制, 避免了网络深度增加后的信息传输过程中的丢失, 使得信息传递更加平滑快速, 因

此也验证了本文提出的密集残差块的有效性和通用性。

## 5 结论

本文提出了一种基于密集残差和双通道注意力机制的前视声呐图像去噪方法。采用了双阶段去噪网络思想, 首先通过双通道注意力机制模块, 捕获全局信息生成图像噪声图, 最后使用密集残差块增强信息提取能力和传递效果生成干净图像。根据前视声呐图像



特性, 使用 FieldII 构建了模拟前视声呐数据集, 模拟了前视声呐噪声分布情况并生成图像噪声对, 利用生成的数据对训练本文算法。在对比实验中, 使用了全参考图像质量评估标准和无参考图像质量评估标准对去噪结果进行分析与评估。在真实数据集和模拟数据集上的实验结果证明了本文提出的算法能够有效抑制前视声呐图像的噪声, 保留图像细节信息。通过模块验证实验和消融实验, 验证了本文所提的双通道注意力机制模块和密集残差模块的有效性和通用性。

**利益冲突:** 所有作者声明无利益冲突

## 参考文献

- [1] Zeng T, Ren L L, Wang Y J, et al. Feature-based underwater three-dimensional sonar target detection and tracking algorithm[J]. *Acta Armam*. <https://doi.org/10.12382/bgxb.2022.0017>.  
曾腾, 任露露, 王宇杰, 等. 基于组合特征的水下三维目标检测跟踪算法[J]. *兵工学报*. <https://doi.org/10.12382/bgxb.2022.0017>.
- [2] Chuang K S, Tzeng H L, Chen S R, et al. Fuzzy c-means clustering with spatial information for image segmentation[J]. *Comput Med Imag Grap*, 2006, **30**(1): 9–15.
- [3] Yang H Y, Wang X Y, Niu P P, et al. Image denoising using nonsubsampling shearlet transform and twin support vector machines[J]. *Neural Netw*, 2014, **57**: 152–165.
- [4] Shao L, Yan R M, Li X L, et al. From heuristic optimization to dictionary learning: a review and comprehensive comparison of image denoising algorithms[J]. *IEEE Trans Cybern*, 2014, **44**(7): 1001–1013.
- [5] Buades A, Coll B, Morel J M. A review of image denoising algorithms, with a new one[J]. *Multiscale Model Simul*, 2005, **4**(2): 490–530.
- [6] Wu D, Du X, Wang K Y. An effective approach for underwater sonar image denoising based on sparse representation [C]//Proceedings of the 2018 IEEE 3rd International Conference on Image, Vision and Computing. Chongqing: IEEE, 2018: 389–393. <https://doi.org/10.1109/ICIVC.2018.8492877>.
- [7] Jiang S Y, Xing C X, Wan Z L, et al. Research on multiplicative speckle noise denoising method of side-scan sonar image based on analysis sparse decomposition[C]//Proceedings of 2021 OES China Ocean Acoustics. Harbin: IEEE, 2021: 1016–1020. <https://doi.org/10.1109/COA50123.2021.9519941>.
- [8] Chen M, Li L, Li Z J, et al. Research on sonar image denoising method based on fixed water area noise model[C]//Proceedings of 2021 IEEE International Conference on Mechatronics and Automation. Takamatsu: IEEE, 2021: 231–235. <https://doi.org/10.1109/ICMA52036.2021.9512575>.
- [9] Stolojescu-Crisan C, Isar A. A new automatic conditioning algorithm for SONAR images[C]//Proceedings of 2021 International Symposium on Signals, Circuits and Systems. Iasi: IEEE, 2021: 1–4. <https://doi.org/10.1109/ISSCS52333.2021.9497424>.
- [10] Fan X J, Li Q W, Huang H, et al. Side-scan sonar image despeckling based on block-matching and 3D filtering[J]. *J Image Grap*, 2012, **17**(1): 68–74.
- [11] 范习健, 李庆武, 黄河, 等. 侧扫声呐图像的3维块匹配降噪方法[J]. *中国图象图形学报*, 2012, **17**(1): 68–74.
- [11] Chen P, Cai X W, Zhao D D, et al. Despeckling for side-scan sonar images based on adaptive block-matching and 3D filtering[J]. *Opto-Electron Eng*, 2020, **47**(7): 190580.  
陈朋, 蔡焯伟, 赵冬冬, 等. 基于自适应BM3D的侧扫声呐图像散斑降噪[J]. *光电工程*, 2020, **47**(7): 190580.
- [12] Zhang K, Zuo W M, Zhang L. FFDNet: toward a fast and flexible solution for CNN-based image denoising[J]. *IEEE Trans Image Process*, 2018, **27**(9): 4608–4622.
- [13] Zhang K, Zuo W M, Chen Y J, et al. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising[J]. *IEEE Trans Image Process*, 2017, **26**(7): 3142–3155.
- [14] Gong X C, Li Z M. An image denoising method using deep asymmetrical skip connection[J]. *J Comput Aided Des Comput Grap*, 2019, **31**(2): 295–302.  
公绪超, 李宗民. 深度非对称跳跃连接的图像降噪方法[J]. *计算机辅助设计与图形学报*, 2019, **31**(2): 295–302.
- [15] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
- [16] Gao S H, Cheng M M, Zhao K, et al. Res2Net: a new multi-scale backbone architecture[J]. *IEEE Trans Pattern Anal Mach Intell*, 2021, **43**(2): 652–662.
- [17] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>.
- [18] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 105–114. <https://doi.org/10.1109/CVPR.2017.19>.
- [19] Wang X N, Huang Y R, Kuang C F, et al. Image restoration of mobile phone under-screen imaging based on deconvolution[J]. *Opto-Electron Eng*, 2021, **48**(6): 200423.  
王晓娜, 黄宇然, 匡翠方, 等. 基于反卷积的手机屏下成像图像恢复[J]. *光电工程*, 2021, **48**(6): 200423.
- [20] Cheng S, Wang Y Z, Huang H B, et al. NBNet: noise basis learning for image denoising with subspace projection [C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 4894–4904. <https://doi.org/10.1109/CVPR46437.2021.00486>.
- [21] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [22] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. *IEEE Trans Pattern Anal Mach Intell*, 2020, **42**(8): 2011–2023.
- [23] Wang Q L, Wu B G, Zhu P F, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11531–11539. <https://doi.org/10.1109/CVPR42600.2020.01155>.
- [24] Li X, Wang W H, Hu X L, et al. Selective kernel

- networks[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 510–519. <https://doi.org/10.1109/CVPR.2019.00060>.
- [25] Wang H L, Yu Q, Li T, et al. Research of software defect prediction model based on CS-ANN[J]. *Appl Res Comput*, 2017, **43**(2): 467–472,476.  
王海林, 于倩, 李彤, 等. 基于CS-ANN的软件缺陷预测模型研究[J]. *计算机应用研究*, 2017, **43**(2): 467–472,476.
- [26] Tian C W, Xu Y, Li Z Y, et al. Attention-guided CNN for image denoising[J]. *Neural Netw*, 2020, **124**: 117–129.
- [27] Guo S, Yan Z F, Zhang K, et al. Toward convolutional blind denoising of real photographs[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 1712–1722. <https://doi.org/10.1109/CVPR.2019.00181>.
- [28] Jensen J A, Svendsen N B. Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers[J]. *IEEE Trans Ultrason Ferroelectr Freq Control*, 1992, **39**(2): 262–267.
- [29] Yu Y J, Acton S T. Speckle reducing anisotropic diffusion[J]. *IEEE Trans Image Process*, 2002, **11**(11): 1260–1270.
- [30] Karabchevsky S, Kahana D, Ben-Harush O, et al. FPGA-based adaptive speckle suppression filter for underwater imaging sonar[J]. *IEEE J Oceanic Eng*, 2011, **36**(4): 646–657.
- [31] Bamber J C, Dickinson R J. Ultrasonic B-scanning: a computer simulation[J]. *Phys Med Biol*, 1980, **25**(3): 463–479.
- [32] Short J R. High-frequency ambient noise and its impact on underwater tracking ranges[J]. *IEEE J Oceanic Eng*, 2005, **30**(2): 267–274.
- [33] Abd-Elmoniem K Z, Youssef A B M, Kadah Y M. Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion[J]. *IEEE Trans Biomed Eng*, 2022, **49**(9): 997–1014.
- [34] Giacomo G, Machado M, Drews P, et al. Sonar-to-satellite translation using deep learning[C]//Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications. Orlando: IEEE, 2018: 454–459. <https://doi.org/10.1109/ICMLA.2018.00074>.
- [35] Mittal A, Moorthy A K, Bovik A C. No-reference image quality assessment in the spatial domain[J]. *IEEE Trans Image Process*, 2012, **21**(12): 4695–4708.
- [36] Geng M F, Meng X X, Yu J Y, et al. Content-noise complementary learning for medical image denoising[J]. *IEEE Trans Med Imaging*, 2022, **41**(2): 407–419.
- [37] Ren C, He X H, Wang C C, et al. Adaptive consistency prior based deep network for image denoising[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 8592–8602. <https://doi.org/10.1109/CVPR46437.2021.00849>.

## 作者简介



赵冬冬(1990-), 男, 博士, 讲师, 博士生导师, 主要研究方向为图像处理、信号处理。

E-mail: zhaodd@zjut.edu.cn



【通信作者】陈朋(1981-), 男, 博士, 教授, 博士生导师, 论文通信作者, 主要研究方向为图像处理、模式识别、嵌入式系统设计。

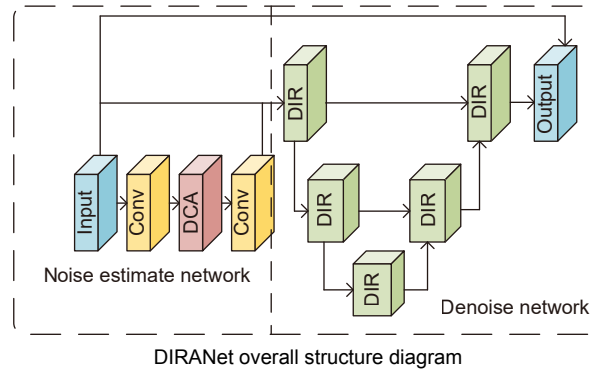
E-mail: chenpeng@zjut.edu.cn



扫描二维码, 获取PDF全文

# Sonar image denoising method based on residual and attention network

Zhao Dongdong<sup>1</sup>, Ye Yifei<sup>1</sup>, Chen Peng<sup>1\*</sup>, Liang Ronghua<sup>1</sup>, Cai Tiancheng<sup>1</sup>, Guo Xinxin<sup>2</sup>



**Overview:** As a kind of underwater active sonar equipment, forward-looking sonar is often used to collect underwater image data. However, it will be affected by underwater noise, which leads degradation of image quality. Due to this situation, a forward-looking sonar image denoising method is proposed based on dense residuals and a dual-channel attention mechanism network. Firstly, the dual attention mechanism is adopted to extract the channel information of the sonar image, collect the global information of the sonar image, and output the noise map of the sonar image. Based on the noise image and sonar image, the dense residual block fully learns the feature information at different scales, and outputs a clean sonar image after multiple learning and information transfer. The main contributions of this paper are as follows.

a) The dual channel attention module is used to estimate the noise, which adaptively accepts information of different scales through two paths using  $3 \times 3$  and  $5 \times 5$  convolutional kernels, respectively, and shunts these information to the next layer of neurons to enhance the feature extraction capability of the module, and then uses one-dimensional convolution to generate a channel attention map to extract the interdependencies between features maps, acquiring more information while reducing the computational volume.

b) Dense in residual module is used to remove noise. This module replaces the  $3 \times 3$  convolutional layers in the traditional residual block with component group convolutions to reduce the number of parameters and improve the training speed, followed by dense and residual connections designed inside the convolutional kernel to learn the differences and connections between feature maps of different sizes, to transfer the learned information from each layer more smoothly and avoid performance degradation and the problem of gradient explosion and disappearance, and finally improve the multi-scale information extraction capability of the network at a fine granularity level.

In this paper, we simulate the generation of forward-looking sonar images by FieldII and add simulated multiplicative Rayleigh noise and additive Gaussian noise to generate a training set for training the network. Later, comparison experiments are conducted on the simulated data test set and the real data set, and the good performance of the proposed method is demonstrated in terms of PSNR, SSIM, and Brisquet image comparison metrics.

Zhao D D, Ye Y F, Chen P, et al. Sonar image denoising method based on residual and attention network[J]. *Opto-Electron Eng*, 2023, 50(6): 230017; DOI: [10.12086/oe.2023.230017](https://doi.org/10.12086/oe.2023.230017)

Foundation item: National Natural Science Foundation of China Youth Science Foundation Project (62001418), Zhejiang Provincial Natural Science Foundation Project (LQ21F010011), Strategic Pioneering Science and Technology Special Project of Chinese Academy of Sciences (Class A) (XDA22030302), and Basic scientific research business expense special fund projects of Zhejiang universities (RF-C2019001)

<sup>1</sup>College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, Zhejiang 330063, China; <sup>2</sup>Institute of Deep-sea Science and Engineering, Chinese Academy of Sciences, Sanya, Hainan 572000, China

\* E-mail: chenpeng@zjut.edu.cn