



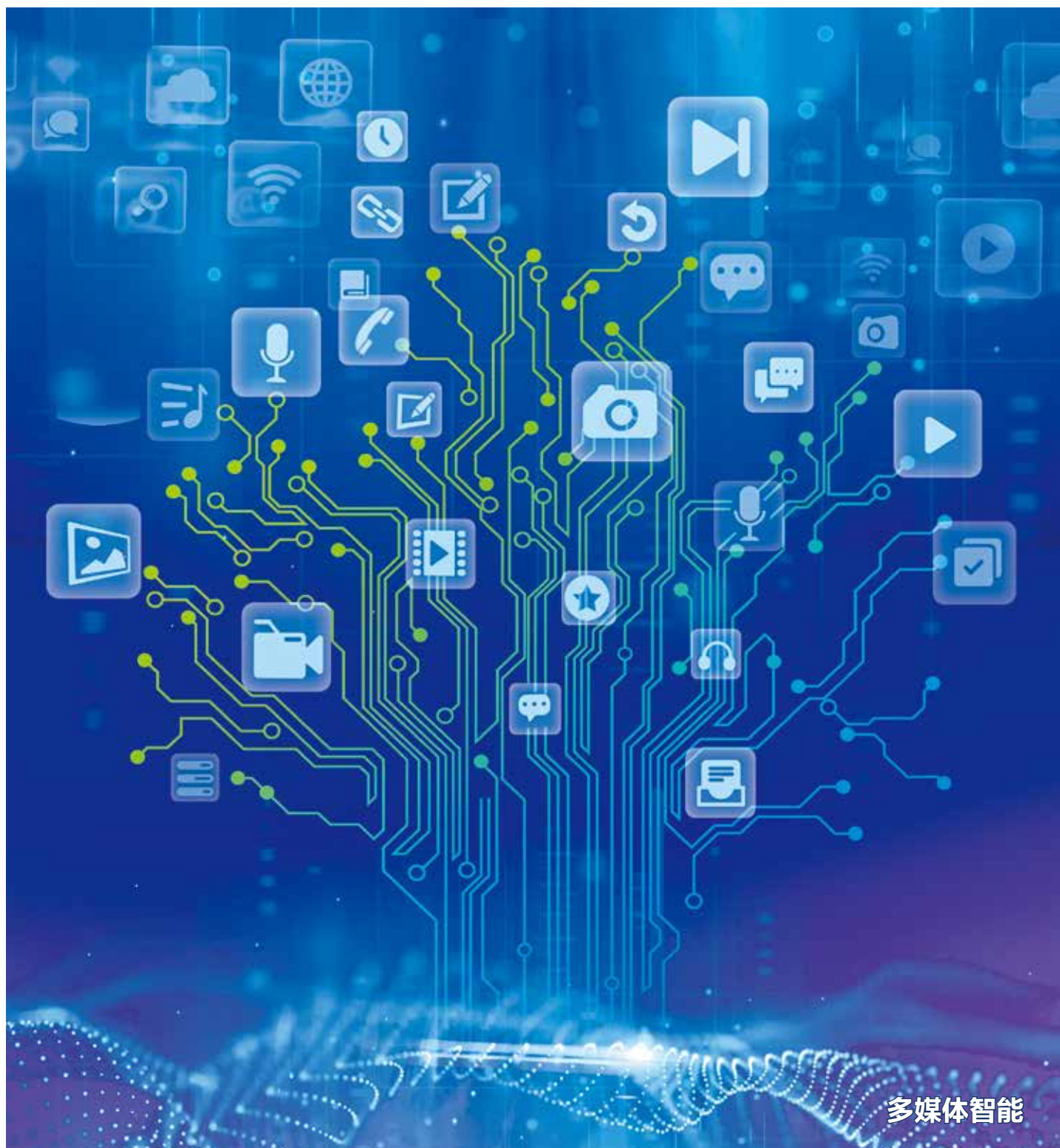
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院  
中国图象图形学学会  
北京应用物理与计算数学研究所

# 中国图象图形学报

2022  
09  
VOL.27

ISSN1006-8961  
CN11-3758/TB



多媒体智能



第27卷第9期（总第317期）  
2022年9月16日

中国精品科技期刊  
中国国际影响力优秀学术期刊  
中国科技核心期刊  
中文核心期刊

## 版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

## Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

**主管单位** 中国科学院  
**主办单位** 中国科学院空天信息创新研究院  
中国图象图形学学会  
北京应用物理与计算数学研究所

**主 编** 吴一戎  
**编辑出版** 《中国图象图形学报》编辑出版委员会  
**通信地址** 北京市海淀区北四环西路19号  
**邮 编** 100190  
**电子信箱** jig@aircas.ac.cn  
**电 话** 010-58887035  
**网 址** www.cjig.cn

**广告发布登记号** 京朝工商广登字20170218号  
**总 发 行** 北京报刊发行局  
**订 购** 全国各地邮局  
**海外发行** 中国国际图书贸易集团有限公司  
(邮政信箱: 北京399信箱 邮编: 100048)  
**印刷装订** 北京科信印刷有限公司

## Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

**Superintended by** Chinese Academy of Sciences  
**Sponsored by** Aerospace Information Research Institute, CAS  
China Society of Image and Graphics  
Institute of Applied Physics and Computational Mathematics

**Editor-in-Chief** Wu Yirong  
**Editor, Publisher** Editorial and Publishing Board of Journal of Image and Graphics  
**Address** No. 19, North 4<sup>th</sup> Ring Road West, Haidian District, Beijing, P. R. China  
**Zip code** 100190  
**E-mail** jig@aircas.ac.cn  
**Telephone** 010-58887035  
**Website** www.cjig.cn

**Distributed by** Beijing Bureau for Distribution of Newspapers and Journals  
**Domestic** All Local Post Offices in China  
**Overseas** China International Book Trading Corporation  
(P.O.Box 399, Beijing 100048, P.R.China)  
**Printed by** Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB  
ISSN 1006-8961  
CODEN ZTTXFZ

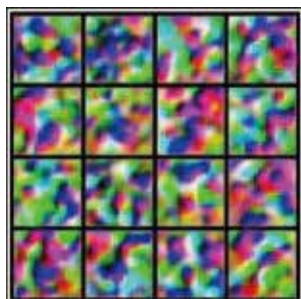
国外发行代号 M1406  
国内邮发代号 82-831  
国内定价 60.00元



多媒体智能：当多媒体遇到人工智能(第2551页)



面向海洋的多模态智能计算：挑战、进展和展望(第2589页)



基于真实数据感知的模型功能窃取攻击(第2721页)

### 《中国图象图形学报》多媒体智能专刊简介

朱文武, 黄庆明, 黄华, 蒋树强, 彭宇新, 刘青山, 王井东, 纪荣嵘, 邓伟洪, 方玉明, 刘家瑛, 韩向娣 ..... 2549

### 学者观点

#### 多媒体智能：当多媒体遇到人工智能

朱文武, 王鑫, 田永鸿, 高文 ..... 2551

#### 视觉知识：跨媒体智能进化的新支点

杨易, 庄越挺, 潘云鹤 ..... 2574

#### 面向海洋的多模态智能计算：挑战、进展和展望

聂婕, 左子杰, 黄磊, 王志刚, 孙正雅, 仲国强, 王鑫, 王玉成, 刘安安, 张弘, 董军宇, 魏志强 ..... 2589

### 综述

#### 基于深度学习的人—物交互关系检测综述

廖越, 李智敏, 刘侃 ..... 2611

#### 人类面部重演方法综述

刘锦, 陈鹏, 王茜, 付晓蒙, 戴娇, 韩冀中 ..... 2629

#### 视觉语言多模态预训练综述

张浩宇, 王天保, 李孟择, 赵洲, 浦世亮, 吴飞 ..... 2652

#### Bayer阵列图像去马赛克算法综述

魏凌云, 孙帮勇 ..... 2683

### 多媒体智能安全

#### 多特征决策融合的音频copy-move篡改检测与定位

张国富, 肖锐, 苏兆品, 廉晨思, 岳峰 ..... 2697

#### 多级特征全局一致性的伪造人脸检测

杨少聪, 王健, 孙运莲, 唐金辉 ..... 2708

#### 基于真实数据感知的模型功能窃取攻击

李延铭, 李长升, 余佳奇, 袁野, 王国仁 ..... 2721

### 目标智能检测

#### 利用时空特征编码的单目标跟踪网络

王蒙蒙, 杨小倩, 刘勇 ..... 2733

#### 结合时空一致性的FairMOT跟踪算法优化

彭嘉淇, 王涛, 陈柯安, 林巍峭 ..... 2749

### 多媒体分析与理解

#### 融合知识表征的多模态Transformer场景文本视觉问答方法

余宙, 俞俊, 朱俊杰, 匡振中 ..... 2761

#### 结合多层级解码器和动态融合机制的图像描述

姜文晖, 占锟, 程一波, 夏雪, 方玉明 ..... 2775

#### 面向非受控场景的人脸图像正面化重建

辛经纬, 魏子凯, 王楠楠, 李洁, 高新波 ..... 2788



# CONTENTS

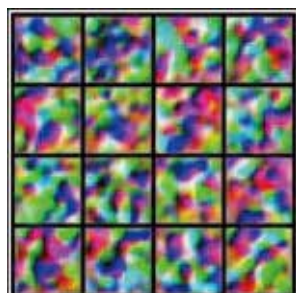
## JOURNAL OF IMAGE AND GRAPHICS



Multimedia intelligence: the convergence of multimedia and artificial intelligence(P2551)



Marine oriented multimodal intelligent computing: challenges, progress and prospects(P2589)



Model functionality stealing attacks based on real data awareness(P2721)

### Scholar View

Multimedia intelligence: the convergence of multimedia and artificial intelligence

Zhu Wenwu, Wang Xin, Tian Yonghong, Gao Wen ..... 2551

The review of visual knowledge: a new pivot for cross-media intelligence evolution

Yang Yi, Zhuang Yueting, Pan Yunhe ..... 2574

Marine oriented multimodal intelligent computing: challenges, progress and prospects

Nie Jie, Zuo Zijie, Huang Lei, Wang Zhigang, Sun Zhengya, Zhong Guoqiang, Wang Xin, Wang Yucheng, Liu An'an, Zhang Hong, Dong Junyu, Wei Zhiqiang ..... 2589

### Review

A review of deep learning based human-object interaction detection

Liao Yue, Li Zhimin, Liu Si ..... 2611

Critical review of human face reenactment methods

Liu Jin, Chen Peng, Wang Xi, Fu Xiaomeng, Dai Jiao, Han Jizhong ..... 2629

Comprehensive review of visual-language-oriented multimodal pre-training methods

Zhang Haoyu, Wang Tianbao, Li Mengze, Zhao Zhou, Pu Shiliang, Wu Fei ..... 2652

The review of demosaicing methods for Bayer color filter array image

Wei Lingyun, Sun Bangyong ..... 2683

### Multimedia Intelligent Security

Multi-feature decision fused detection and localization method for copy-move forgery of digital audio clips

Zhang Guofu, Xiao Rui, Su Zhaopin, Lian Chensi, Yue Feng ..... 2697

Multi-level features global consistency for human facial deepfake detection

Yang Shaocong, Wang Jian, Sun Yunlian, Tang Jinhui ..... 2708

Model functionality stealing attacks based on real data awareness

Li Yanming, Li Changsheng, Yu Jiaqi, Yuan Ye, Wang Guoren ..... 2721

### Object Intelligent Detection

A spatio-temporal encoded network for single object tracking

Wang Mengmeng, Yang Xiaoqian, Liu Yong ..... 2733

Spatio-temporal consistency based FairMOT tracking algorithm optimization

Peng Jiaqi, Wang Tao, Chen Kean, Lin Weiyao ..... 2749

### Multimedia Analysis and Understanding

Knowledge-representation-enhanced multimodal Transformer for scene text visual question answering

Yu Zhou, Yu Jun, Zhu Junjie, Kuang Zhenzhong ..... 2761

The integrated mechanism of hierarchical decoders and dynamic fusion for image captioning

Jiang Wenhui, Zhan Kun, Cheng Yibo, Xia Xue, Fang Yuming ..... 2775

Face frontalization for uncontrolled scenes

Xin Jingwei, Wei Zikai, Wang Nannan, Li Jie, Gao Xinbo ..... 2788



中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2022)09-2683-14

论文引用格式: Wei L Y and Sun B Y. 2022. The review of demosaicing methods for Bayer color filter array image. Journal of Image and Graphics, 27(09): 2683-2696 (魏凌云, 孙帮勇. 2022. Bayer 阵列图像去马赛克算法综述. 中国图象图形学报, 27(09): 2683-2696) [DOI: 10.11834/jig.211059]

# Bayer 阵列图像去马赛克算法综述

魏凌云<sup>1</sup>, 孙帮勇<sup>1, 2\*</sup>

1. 西安理工大学印刷包装与数字媒体学院, 西安 710048;

2. 中国科学院西安光学精密机械研究所光谱成像技术重点实验室, 西安 710119

**摘要:** Bayer 阵列图像去马赛克技术是对稀疏采样的 Bayer 阵列图像进行 RGB 信息重建, 图像重建质量是成像设备评价的重要因素之一, 同时也对其他计算机视觉任务 (如图像分割、人脸识别) 产生影响。随着深度学习方法的快速发展, 图像去马赛克领域提出了多种高性能算法。为了便于研究者更全面地了解图像去马赛克算法的原理和研究进展, 本文对该领域的经典算法和深度学习算法进行综述。首先对 Bayer 采样阵列原理和图像去马赛克技术进行概述。然后将现有方法分为传统方法和基于深度学习方法两类进行总结, 同时根据去马赛克任务是否具有独立性, 将深度学习方法分为独立去马赛克任务和联合去马赛克任务两类, 分析不同方法的原理和优缺点, 重点阐述基于深度学习的去马赛克方法的网络结构和重建机理, 介绍去马赛克领域常用的公共数据集和性能评价指标, 并对图像去马赛克相关实验进行分析对比。最后, 围绕网络深度、运算效率和实用性等方面分析了现阶段图像去马赛克技术面临的挑战及未来发展方向。目前, 基于深度学习的图像去马赛克方法已成为主流发展方向, 但仍然存在计算成本较高、实际应用性不强等问题。因此, 如何开发出重建精度高、处理时间短以及实用性强的图像去马赛克方法, 是该领域未来重要的研究方向。

**关键词:** 图像去马赛克; Bayer 阵列图像; 图像处理; 深度学习; 卷积神经网络 (CNN); 综述

## The review of demosaicing methods for Bayer color filter array image

Wei Lingyun<sup>1</sup>, Sun Bangyong<sup>1, 2\*</sup>

1. College of Printing, Packaging and Digital Media, Xi'an University of Technology, Xi'an 710048, China;

2. Key Laboratory of Spectral Imaging Technology, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China

**Abstract:** The Bayer array image demosaicing technology is one of the essential parts in the digital image signal processing process mechanism, which reconstructs the sparsely sampled Bayer array image into a completed RGB image. In terms of a single-sensor image acquisition of Bayer color filter array, single color channel value is just recorded by the position of each pixel. The other two color channels information needs to be reconstructed by the Bayer color filter array demosaicking technology. The reconstruction effect of qualified RGB information determines the accuracy of computer vision analysis like image segmentation and face recognition. Deep learning based Bayer array image demosaicing technology has promoted more

收稿日期: 2021-11-24; 修回日期: 2022-05-07; 预印本日期: 2022-05-14

\* 通信作者: 孙帮勇 sunbangyong@xaut.edu.cn

**基金项目:** 国家自然科学基金项目 (62076199); 陕西省重点研发计划资助 (2021GY-027, 2022ZDLGY01-03); 中国科学院光谱成像技术重点实验室基金项目 (LSIT201801D)

**Supported by:** National Natural Science Foundation of China (62076199); Key R&D Project of Shaan'xi Province, China (2021GY-027, 2022ZDLGY01-03); Key Laboratory of Spectral Imaging Technology of Chinese Academy of Sciences (LSIT201801D)

high-performance algorithms nowadays. To recognize the principle and development of image demosaicing algorithms, we review some classic algorithms and analyze the key concepts. First, an overview is illustrated for the Bayer sampling array and image demosaicing technology. Next, we divide the existing methods into two categories in accordance with traditional methods and deep learning methods. The pros and cons of different methods are analyzed and the information transmission issue of the learning-based demosaicing network is focused on. We introduce traditional demosaicing methods in brief based on the basic principle, growth path and applications of the residual interpolation algorithms. Specifically, we divide the learning-based methods into single demosaic algorithms and joint multi-task demosaic algorithms derived of the independence of the demosaicing task. Single demosaicing algorithms can be divided into three categories of those are two-stage demosaicing network, three-stage demosaicing network, and end-to-end demosaicing network. Joint multi-task demosaic algorithms are mostly based on end-to-end multiple networks structure like full convolution, residual, densely connected, U-Net, feature pyramid, and generative adversarial network (GAN). At the same time, we compare seven representative methods to the two benched datasets of Kodak and McMaster, including three traditional methods and four deep learning methods. In addition, we introduce common public datasets and performance evaluation metrics peak signal-to-noise ratio (PSNR) and structural similarity in the field of demosaicing. Finally, we analyze the technical problems faced by the current image demosaicing technology and the future development direction from the perspectives of its network depth, computing efficiency, and feasibility. Our experimental results show that the deep learning based method surpasses the traditional methods in terms of objective quality evaluation indicators and subjective vision. We clarified that the highest PSNR value of the learning-based methods on the Kodak dataset surpassed the traditional methods by 12.73 dB, and the highest PSNR value of the learning-based methods on the McMaster dataset surpassed the traditional methods by 5.10 dB. The traditional method obtains the estimated value of the color image via the degradation model of the Bayer array image and some prior information. However, some distortion issues appear in the clear edges and multi-textured areas of the reconstructed image, such as color artifacts and checkerboard effects. Current deep learning methods have become the main development prospects in the field of demosaicing. Although the reconstructed image of learning-based methods is more similar to the original image than the traditional method on the visual effect and accuracy, it also has the problem of high computational cost and poor practical applicability. It is challenged to guarantee the accuracy of network algorithms and reduce the computational cost of the network both. Therefore, a high-performance and effective image demosaicing method has to be developed further in terms of datasets option, network structure optimization, and improvement of computing efficiency.

**Key words:** demosaicking; Bayer array image; image processing; deep learning; convolutional neural network (CNN); review

## 0 引言

滤色片、棱镜和光栅是直接成像最常用的分色元件,由于滤色片可直接覆盖在传感器表面对入射光进行分色,能极大降低成像设备的体积和成本,因此彩色数字相机多采用滤色片作为分光元件。同时,为实现单次曝光成像,数字相机大多设计为单个传感器表面覆盖不同颜色滤色片的结构,通过稀疏采样和实时插值重建,实现图像的快速获取。

数字相机直接成像获取的图像为彩色滤色片阵列(color filter array, CFA)图像,每个像素位置仅存储一种特定颜色的灰度值,因此这种阵列图像也称做马赛克图像。对于 RGB 彩色相机来说,马赛克图

像的采样率为  $1/3$ ,每个位置上另外两种未采样颜色需通过重建算法获得,这个过程称为图像去马赛克。最通用的阵列图像为 Bayer CFA (Bayer, 1976),如图 1(a)所示,其在空间分布上绿色采样量是红色或蓝色的两倍。此外,在 RGB 成像领域,为了达到较好的重建效果,研究者设计出其他 CFA: X-Trans CFA (图 1(b)) 和 RGBW CFA (RGB-white CFA) (图 1(c)),并提出相应的去马赛克方法 (Rafinazari 和 Dubois, 2014; Kang 和 Jung, 2020; 王海琳 等, 2021)。

针对欠采样条件下的 Bayer 阵列图像,利用去马赛克数学模型重建未采样颜色值,实现完整图像信息重建,是 RGB 相机成像中最普遍的方法。其能够显著缓解数字相机的硬件成本,并极大提高图像获取的便利性。然而,由于采样率仅为  $1/3$ ,多数

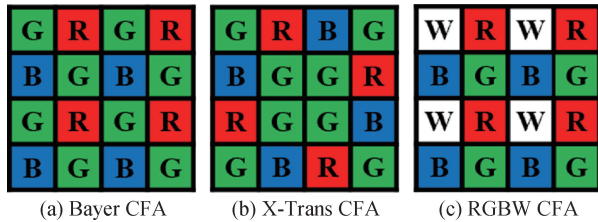


图1 RGB 成像领域中的多种 CFA

Fig.1 Various CFAs in RGB imaging

((a) Bayer CFA; (b) X-Trans CFA; (c) RGBW CFA)

图像去马赛克方法存在龟纹、拉链效应和颜色失真等缺陷,因此去马赛克长期以来一直是成像和视觉领域的挑战性问题 and 研究热点之一。研究发现,去马赛克方法可大致分为传统方法和深度学习方法两类。传统方法主要依赖特定的先验知识或人工设计的模型,按照一定规则对各像素颜色值进行重建,泛化能力较差。深度学习方法则是根据大量的数据集训练一种图像重建网络,学习 Bayer 阵列图像到原始完整图像之间的映射关系,所得模型具有更高的重建效率和泛化能力。本文主要对 Bayer CFA 去马赛克算法进行回顾和分析,着重对深度学习方法重建模型进行总结和分类,并按照独立去马赛克任务和联合其他视觉任务两个方向进行阐述。

## 1 传统方法

前期去马赛克算法是较为简单的色差插值方法,如最近邻插值(Adams, 1995; Ramanath 等, 2002)、双线性插值(Hou 和 Andrews, 1978)、双三次插值(Li 和 Randhawa, 2007)和三次 B 样条插值(Longère 等, 2002)等。色差插值法简单且容易实现,但忽略了通道间的相关性和边缘结构细节,重建

结果经常存在颜色伪影、拉链效应和模糊等缺陷,仅适合重建较为平滑的图像类型。后来, Pekkuksen 和 Altunbasak (2010) 对色差法进行改进,提出基于梯度的无阈值算法 (gradient based threshold free, GBTF), 利用色差梯度计算每个方向的权重, 然后根据每个方向的权重, 结合来自各个方向的估计对图像值进行插值。GBTF 虽然在客观评价方面有一定程度的提高,但在主观评价中仍存在颜色伪影、模糊等失真现象。

为降低传统色差插值误差, Kiku 等人 (2013) 验证了残差插值 (residual interpolation, RI) 具有更小的误差, 能够产生较好的重建效果, 从而提出基于残差插值的去马赛克方法。该算法以 GBTF 方法的框架为基础, 将残差插值代替色差插值, 实现马赛克图像的重建。其中, 残差是指最初采样到的像素值与初步估计值之间的差值。如图 2 所示, 以恢复 R 像素为例, 首先对 G 像素进行插值, 然后使用引导滤波 (He 等, 2013) 对初步采样到的 R 像素值上采样, 得到初步估计的 R 值  $\hat{R}$ , 接着计算残差  $R - \hat{R}$ , 并对残差进行双线性插值, 得到估计的残差平面, 最后通过残差平面和初步估计值  $\hat{R}$  相加得到最终 R 通道图。相比于 GBTF, 虽然 RI 提升的效果不大, 但为后面出现的一系列残差算法奠定了基础。例如, Kiku 等人 (2014) 进一步对 RI 进行改进, 提出最小化拉普拉斯残差插值方法 (minimized-Laplacian residual interpolation, MLRI), 指出双线性插值可以为具有较小拉普拉斯能量的图像提供更好的插值结果。在计算流程方面, 与 RI 相比, MLRI 得到初步估计 R 通道像素值  $\hat{R}$  后, 通过最小化残差的拉普拉斯能量计算出需要插值的残差, 然后再与初步估计的像素值相加进行重建。

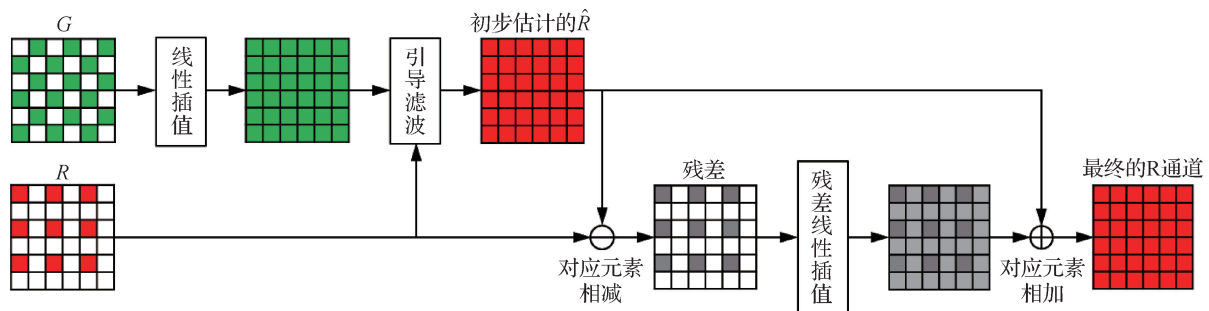


图2 残差插值计算流程

Fig.2 Workflow of residual interpolation



以上基于 RI 的重建方法中,由于只能恢复 R 和 B 通道的像素值,而对于 G 通道仍是运用色差插值方法,从而不能充分挖掘 G 通道信息,并且 G 通道重建中引入的误差将会影响 R 和 B 通道的重建。针对此问题,Ye 和 Ma (2015) 提出迭代残差方法 (iterative residual interpolation, IRI),通过迭代运算对三通道应用 RI,使通道间相互引导重建。之后,Monno 等人(2015)提出自适应残差插值方法 (adaptive residual interpolation, ARI),不同于 IRI (IRI 对整个图像像素使用相同的迭代次数),ARI 自适应地为每个像素选择合适的迭代次数,使迭代次数更灵活、更适合每个通道像素值的重建。

## 2 深度学习方法

由于深度学习中的卷积神经网络 (convolutional neural networks, CNN) 能够有效提取图像的浅层、深层特征以及结构信息,已经在图像超分辨率 (Liu 等, 2020a; 雷鹏程 等, 2020)、去噪 (吴从中等, 2018; Jia 等, 2021) 和去模糊 (吴迪 等, 2020; Wan

等, 2021) 等视觉处理任务中取得令人满意的结果。近年来,基于深度学习的去马赛克方法发展较快,其利用大量数据集进行网络训练,不断学习马赛克图像与原始图像之间的映射关系,从而获得精度较高、泛化能力较强的图像重建模型。研究发现,与传统方法相比,基于深度学习的去马赛克方法重建图像质量得到进一步提升,在主观评价和客观评价指标上都更接近原始图像,在多种应用领域已逐步替代传统方法。

如图 3 所示,按照去马赛克任务的独立性,基于深度学习的方法可大致分为独立去马赛克任务和联合其他视觉任务的去马赛克任务。独立去马赛克方法中,有些方法首先建立参数学习的深度网络,然后利用学习到的参数在传统去马赛克模型下进行图像重建;有些方法是端到端网络,即针对给定的马赛克图像直接输出重建图像。联合去马赛克任务,主要考虑 Bayer 阵列图像的采样率低、噪声明显等特征,因此已提出了大量去噪和去马赛克联合方法,以及图像超分辨率和去马赛克的联合处理方法。

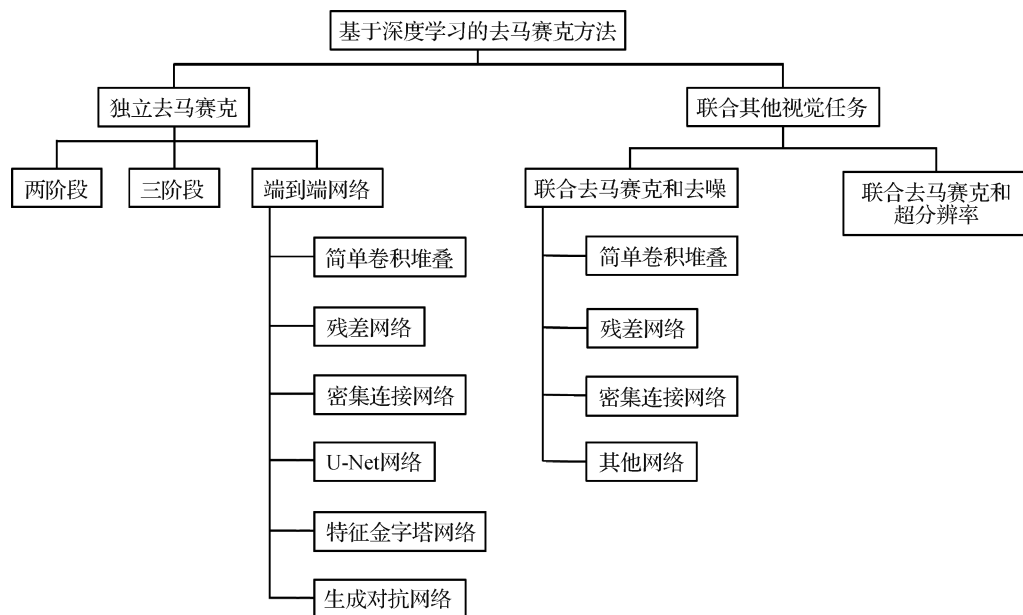


图3 基于深度学习的去马赛克方法分类

Fig. 3 The classification of deep learning demosaicing method

### 2.1 独立去马赛克任务

独立去马赛克任务可大致分为两阶段去马赛克、三阶段去马赛克和端到端去马赛克3类。其中两阶段去马赛克和三阶段去马赛克方法一般遵循传

统去马赛克流程,先恢复采样率较高的 G 通道,然后以此作为先验信息,引导 R、B 两通道的像素信息重建。考虑到端到端卷积神经网络重建效率高、泛化能力强的特点,近年来更多的图像去马赛克网络

设计成端到端方式。

### 2.1.1 两阶段去马赛克

Tan 等人(2017)提出 CDM-CNN(color demosaicking-CNN), 首先用双线性插值生成 Bayer 阵列图像中的缺失像素值, 然后构建图像去马赛克 CNN 来恢复 G 通道, 并将恢复的 G 通道作为引导图像来重建 R 和 B 通道。随后, 余继辉等人(2020)提出插值预处理后利用网络进行重建的两阶段方法。Tan 等人(2020)根据 Bayer CFA 阵列特点, 采用不同的特征提取网络获得 R、G、B 三通道的浅层特征后进行特征融合, 从而得到重建的 R、G、B 单通道图, 然后进行特征增强合并成 RGB 全彩色图, 完成图像重建。Tan 等人(2018)根据不同图像的复杂程度, 利用梯度校正双线性插值方法对 Bayer 阵列图像插值后, 建立全卷积网络进行图像重建, 该网络针对一般图像、平滑图像和粗糙图像训练了 3 类网络参数, 3 个网络生成的图像进一步融合后作为最终的去马赛克图像。

两阶段去马赛克方法继续保留了传统去马赛克的流程, 通过构建卷积神经网络进行关键参数的学习, 与传统方法相比, 性能得到一定提升, 峰值信噪比(peak signal to noise ratio, PSNR)等客观指标提高明显。同时, 部分两阶段去马赛克方法根据 Bayer 阵列的点阵结构和重建图像的复杂程度, 设计了不同的网络进行参数学习, 使模型有针对性地进行重建图像, 获得较好的主观评价结果。但是, 已提出的两阶段去马赛克网络相对较浅, 未充分利用图像特征, 仍存在一定提升空间。此外, 大部分两阶段的方法将插值作为初始化, 可使后续网络稳定且易于训练。但值得注意的是, 插值带来的误差也将会成为后续网络学习的对象, 从而影响网络的表达能力。

### 2.1.2 三阶段去马赛克

Cui 等人(2018)分析了 R、G、B 三通道之间的相关性, 指出 R、G、B 三通道两两之间具有不同程度的相关性, 且 G 通道与 R、B 通道具有更高的相关性, 提出三阶段(3-stage)去马赛克方法。该方法首先构建基本卷积神经网络计算 G 通道, 再将重建的 G 图像分别合并到 R、B 通道中进行引导, 重建 R、B 通道, 最后合成完整的 RGB 三通道图像。此方法遵循传统去马赛克流程, 同时运用卷积神经网络有效提取图像特征。随后, Wang 等人(2021b)指出 3-stage 的初始化和先验信息融合不够有效, 为解决

此问题, 在遵循 3-stage 的去马赛克流程的同时, 构建出新的三阶段深度卷积网络(new three-stage deep convolutional network, NTSDCN), 将采用拉普拉斯能量损失函数约束的局部残差学习结构作为新的特征提取方式, 并提出新的引导方式, 在特征域应用非齐次线性模型有效融合 G 图像先验信息。但是该方法网络结构相对较深, 参数数量大, 导致计算成本高, 在实际应用方面存在一定障碍。

Niu 和 Ouyang(2020)针对遵循传统去马赛克流程的深度神经网络学习的不足, 如 3-stage 模型中时间成本高、参数多等问题, 将遵循传统去马赛克流程的深度神经网络分解为三阶段轻量级网络。不同于其他方法通过构建网络计算出图像特征或残差特征之后再融合重建的思路, 该方法是通过构建神经网络估计色差后再进行重建。首先通过高质量线性插值重建出 Bayer 图像中的未采样值, 然后建立卷积神经网络恢复 G 通道并分别估计 R-G 和 B-G 色差, 最后根据色差分别重建 R 和 B, 得到最终的 RGB 图。该方法充分利用了图像的跨通道相关性, 并且可以独立训练和并行估计。Iriyama 等人(2021)对三阶段轻量级网络进行了改进, 将通道间的相关性和自相似性有效结合, 在构建网络估计 R-G 和 G-B 色差的同时, 应用非局部注意模块获取色差域中的长期相关性, 从而可以更准确地预测纹理和边缘区域中的高频分量, 提高重建质量。

此外, Kim 等人(2020)考虑到 Bayer CFA 的点阵结构特点, 为提高网络性能, 使其更有针对性, 提出了一种基于自适应网络感知滤波生成器和全局细化单元的去马赛克网络, 首先通过密集连接网络生成自适应感知滤波器, 然后在  $5 \times 5$  的窗口中对局部信息进行插值, 最后使用全局细化单元利用全局信息进行细化, 得到最终重建的去马赛克图像。

三阶段去马赛克网络具有较高的重建精度, 在遵循传统去马赛克流程的基础上, 针对特征提取方式、引导方式和色差估计方法等方面进行改进, 充分利用通道间的相关性, 提取较为准确的图像特征, 使重建图像能够保留更丰富的图像边缘和细节。然而, 在性能提升的同时, 网络结构也相应加深、内存和时间成本加大, 导致难以投入到实际应用中。

### 2.1.3 端到端卷积神经网络

随着卷积神经网络结构的发展, 多种新的网络

结构应用到去马赛克中,如简单卷积堆叠(Syu 等, 2018)、残差网络(residual neural network, ResNet)(Kokkinos 和 Lefkimmatis, 2018; Verma 等, 2020; Zhang 等, 2019)、密集连接网络(Park 和 Jeong, 2018, 2019)、U-Net 网络(Kang 等, 2019; Wang 等, 2021a)、特征金字塔网络(Mei 等, 2020)和生成对抗网络(generative adversarial network, GAN)(Zhao 等, 2019; Luo 和 Wang, 2020)等,重建性能均有不同程度的提升。然而,近年来越来越多的端到端网络只专注于通过增加网络深度提升重建性能,忽略了参数数量与时间成本的影响,未充分权衡性能与成本,实用性不强。

1)简单卷积堆叠。Syu 等人(2018)分别受图像超分辨率模型 SRCNN(super-resolution CNN)(Dong 等, 2014)和 VDSR(very deep super-resolution)(Kim 等, 2016)的启发,提出了 DMCNN(demosaicing CNN)和 DMCNN-VD(DMCNN-very deep)。如图 4

(a)所示,DMCNN 与 SRCNN 类似,由 3 个卷积层堆叠而成,第 1 层为特征提取层,提取图像浅层特征;第 2 层为非线性映射层,学习 Bayer 阵列图像到完整 RGB 图像之间的非线性映射关系;第 3 层为重建层,直接重建出完整的 RGB 图像。3 个卷积核的尺寸分别为  $9 \times 9 \times 128$ 、 $1 \times 1 \times 64$  和  $5 \times 5 \times 3$ 。尽管 DMCNN 实现了端到端的学习方式,但由于该网络比较浅,没有充分挖掘深度学习模型的潜力。为此, Syu 等人(2018)提出了 DMCNN-VD。如图 4(b)所示,DMCNN-VD 由 20 层构成,每层包含卷积层、批量归一化(Ioffe 和 Szegedy, 2015)和缩放指数型线性单元激活函数(scaled exponential linear units, SELU)(Klambauer 等, 2017)。同时,为了防止梯度消失和梯度爆炸问题,还加入了残差学习策略,达到较快的收敛速度。DMCNN 和 DMCNN-VD 是对端到端独立去马赛克任务的初步尝试,表明了端到端卷积神经网络模型在去马赛克任务中的有效性。

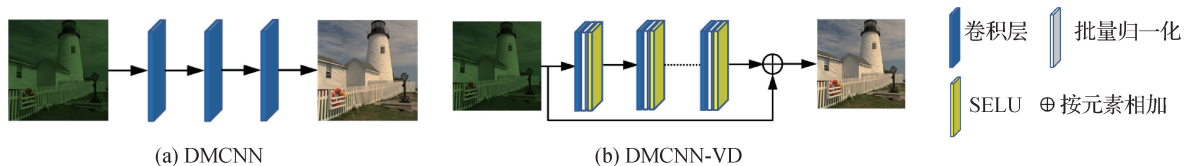


图 4 DMCNN 和 DMCNN-VD 的网络结构

Fig. 4 The network architecture of DMCNN and DMCNN-VD((a)DMCNN;(b)DMCNN-VD)

2)残差网络。尽管网络越深感受野越大,提取的上下文信息越丰富,模型的重建精度越好,但深度网络的参数量大,不易收敛,易出现梯度消失或梯度爆炸现象。为解决网络退化现象、缓解梯度消失和梯度爆炸问题,加快收敛速度,He 等人(2016)提出残差网络结构,如图 5(a)所示。此后,研究者将残差结构引入到去马赛克任务,Kokkinos 和 Lefkimmatis(2018)提出用迭代的方式排列残差块,指出即使在较小的训练数据集上也能训练出较好的模型,网络参数较少,但是网络中残差块数量较少,效果提升不明显。Zhang 等人(2019)提出残差非局部注意力网络

(residual non-local attention network, RNAN),通过考虑像素间的全局相关性来获取非局部特征,进一步增强网络的表达能力。Verma 等人(2020)提出将瓶颈残差(ResNet-bottleneck)网络应用到去马赛克任务中,如图 5(b)所示,网络包括 10 个残差块,残差块的输入和输出都是 256 个通道,里面的特征图通道数设置为 64,并在残差结构的基础上,运用瓶颈结构加深网络,由于参数量较少,更容易训练,使网络具有很好的通用性。然而,该网络未考虑图像信息分布不均匀问题。之后,余继辉和杨晓敏(2021)设计了残差注意力模块,充分利用了特征图之间的相关性和特征图内的空间相关性。

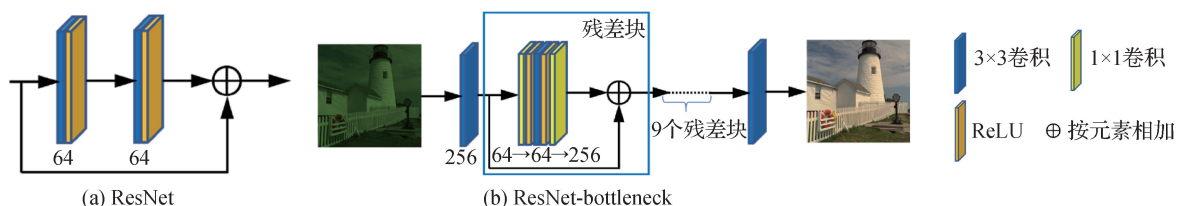


图 5 残差结构和瓶颈残差网络

Fig. 5 The architecture of ResNet and ResNet-bottleneck network((a)ResNet;(b)ResNet-bottleneck)



3) 密集连接网络。由于之前的卷积网络结构是按顺序提取图像特征图,未充分利用到上一步以外的其他层特征信息,因此,Huang 等人(2017)提出密集连接网络,如图 6(a)所示,在网络中,每层卷积层的输入是前面所有层的输出合并,确保网络具有最大的信息流通,减少了参数数量,降低了过拟合问题出现的概率。随后,Park 和 Jeong(2018)提出密集连接去马赛克网络,如图 6(b)所示,网络包括 15 个密集块,每个密集块均由卷积块(前 6 个卷积

层)和过渡块(最后  $1 \times 1$  卷积层)组成,每个卷积层紧密相连,可充分利用每一层的特征信息,有利于图像重建。在网络的最后通过亚像素插值层(Shi 等,2016)生成重建图像。该网络不需要初始插值预处理,同时过渡块可以减少特征数量,计算复杂度不高,更易训练,重建精度更高。之后,Park 和 Jeong(2019)指出密集连接网络忽略了对浅层特征的利用,提出密集连接残差网络,将密集连接网络与残差结构有效结合,进一步提高了网络的重建能力。

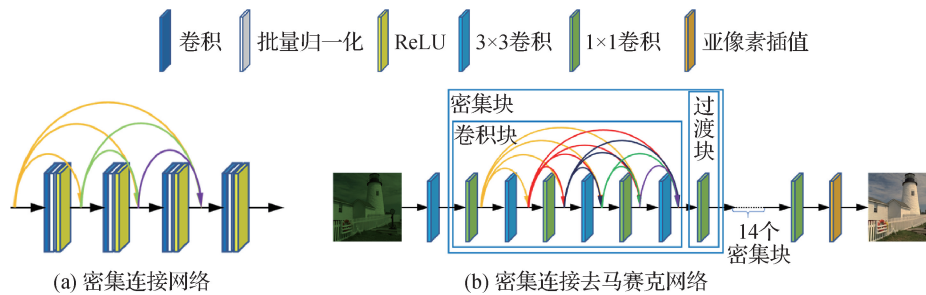


图 6 密集连接网络和密集连接去马赛克网络

Fig. 6 The architecture of densely connected network and densely connected demosaicing network  
( (a) densely connected network; (b) densely connected demosaicing network )

4) U-Net 网络。U-Net 最初由 Ronneberger 等人(2015)提出并应用于图像分割任务,其将多尺度的特征融合在一起,能够处理任意大小的图像,具有较好的灵活性。如图 7(a)所示,U-Net 以编码器—解码器为基础,将原始图像的特征编码提取后,再将特征进行解码还原为需要的信息,结构简单且有效。相应地,针对去马赛克任务,Kang 等人(2019)提出一种基于通道注意力的多尺度多层次特征融合的去马赛克算法,利用 U-Net 结构获取多尺度、多层次的特征,并引入通道注意机制,对每个场景的通道特征进行自适应调整,提高网络的表现力。后来,Zhou 等人(2018b)指出 U-Net 结构(Ronneberger 等人,2015)中位于前面的卷积层感受野较小,提取的是

局部特征,而越往后感受野越大,提取到的特征更接近全局特征。如果将差异较大的局部特征与全局特征直接融合,会增大网络的难度。因此,为了减少语义差别,Zhou 等人(2018b)提出了 U-Net++。如图 7(b)所示,在 U-Net 结构中直接连接的基础之上增加了类似于密集连接结构的卷积层,然后再融合下一阶段的特征,同时配合深度监督策略,使网络可任意剪枝。与 U-Net 相比,U-Net++ 仅增加了 16% 的参数量,而精度得到较大提升。Wang 等人(2021a)对 U-Net++ 进行改进,应用到边缘计算设备去马赛克任务中,其利用不改变图像大小的高斯模糊层代替下采样操作,虽然 PSNR 值不高,但重建图像中的伪影抑制较好,主观评价质量较高。

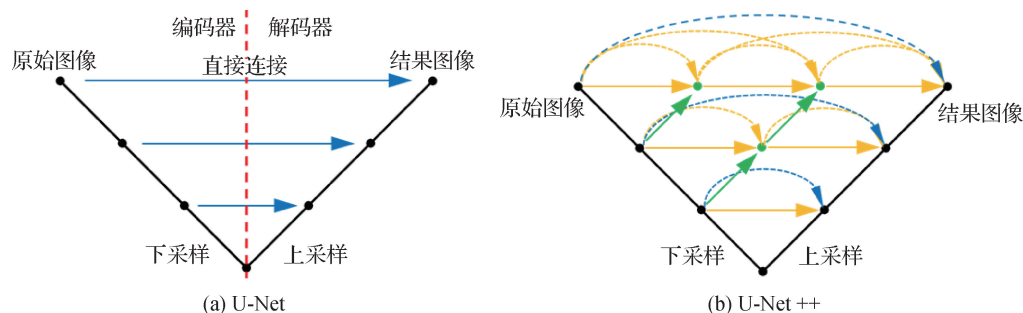


图 7 U-Net 和 U-Net++ 的结构

Fig. 7 The architecture of U-Net and U-Net++ ( (a) U-Net; (b) U-Net++ )



5) 特征金字塔网络。Mei 等人(2020)提出金字塔注意网络,将多尺度特征提取方式与注意力机制结合进行重建,充分利用图像的自相似性来捕获远程特征对应关系。同时,该网络中的金字塔注意模块是一个通用的构建模块,可以灵活地集成到多种网络框架中。

6) 生成对抗网络。生成对抗网络(Goodfellow 等,2014)在图像生成领域发展较快,越来越多的研究者尝试将 GAN 的训练策略运用于图像去马赛克领域,以生成高质量的重建图像。Zhao 等人(2019)将 U-Net 作为生成器,同时将生成显著图作为数据增强的一部分,提高训练集的丰富性。Luo 和 Wang(2020)将 U-Net 作为生成器,将密集残差网络作为鉴别器,进行交替训练,得到网络模型,重建图像可以更接近原始图像。罗静蕊等人(2021)利用生成对抗网络结合网络损失函数增强学习图像高频信息的能力,提升网络整体性能。

## 2.2 联合去马赛克任务

去噪、去马赛克和超分辨率均是彩色成像图像处理流程中的关键步骤。直接对未去噪的 Bayer 阵列图像进行去马赛克处理,重建网络会将噪声信息作为有用信息进行学习,导致重建效果不理想。因此,研究者经常采用去马赛克与去噪联合处理的方法,建立端到端网络对 Bayer 阵列图像实现一次性重建。此类方法方便快捷,可一次性处理多个任务,但针对性不强,精度偏低。

### 2.2.1 去马赛克和去噪联合任务

1) 简单卷积堆叠。Gharbi 等人(2016)采用类似 SRCNN(Dong 等,2014)的网络结构,建立 Bayer 阵列图像与完整信息图像之间的非线性映射关系,首次将卷积神经网络运用于联合去马赛克和去噪任务,相比之前方法,在重建精度上得到一定程度提高。但是,该网络在训练过程中依赖大量数据集,并且设计的网络在输出图像尺寸方面不稳定,实施比较困难。Prakash 等人(2017)也采用简单堆叠卷积层进行去马赛克和去噪,但重建的图像仍存在一定的误差和伪影。

2) 残差网络。Huang 等人(2018)将深度残差学习和聚合残差变换(Xie 等,2017)的概念应用于联合去马赛克和去噪任务,提出较少参数量的轻量级网络,但由于网络包含了 36 个卷积层,其内存读取写入的总时间成本不可忽略。另外,Kokkinos 和

Lefkimmatis(2019)提出以迭代的方式排列残差网络,能够有效利用提取到的特征,提高网络表达能力。Guo 等人(2020)针对成本问题,将 Inception-ResNet 结构(Szegedy 等,2017)进行改进,得到轻量级的 Inception 结构,应用到残差网络中,获得了更好的跨通道融合特征和更大的接受域,能够有效避免棋盘效应且可以保留更多的图像细节。随后,Xing 和 Egiiazarian(2021)在残差网络的基础上引入了通道注意力机制,也取得不错的重建效果。

3) 密集连接网络。Din 等人(2020)提出用密集连接网络先进行去马赛克再去噪,但忽略了去马赛克后的噪声模型破坏问题,导致去噪的结果不够准确。Qian 等人(2021)提出用密集连接残差网络进行去马赛克、去噪和超分辨率的联合任务,其将网络划分为两个模块,先进行图像超分辨率,再执行去马赛克和去噪,可以减少去马赛克伪影,同时使去马赛克任务突破分辨率的限制。

4) 其他网络。Dong 等人(2018)提出一种端到端的基于 GAN 的联合去马赛克去噪网络,并且应用具有感知损失函数和对抗性损失函数的鉴别器网络来优化重建图像的感知质量,使重建图像更接近原始图像。Liu 等人(2020b)指出对所有位置和所有图像不应该用同一组参数,应考虑图像的具体内容和重建难度,因此提出两种新的自引导方法——绿色通道引导和密度图引导方法,使网络可自适应处理图像不同层次的特征,重建精度更准确。Park 等人(2020)提出一种可对单个图案图像进行训练,可适用于不同噪声水平的变分深度图像先验网络。Sharif 等人(2021)运用深度注意力机制和空间注意力机制,引入感知颜色损失和正则化特征损失两个新的感知损失,提高重建图像的感知质量。

### 2.2.2 去马赛克和超分辨率联合任务

去马赛克网络的预处理操作大多是将采样得到的 Bayer 阵列图像下采样成半尺寸的四通道 RGGB 图像,导致分辨率降低或图像细节丢失。同时,RGB 的相对位置信息也无法保证。因此,学者提出将去马赛克和超分辨率进行联合处理,以有效保留图像的颜色信息和位置信息。

Zhang 等人(2018)提出用全局残差结构学习图像特征,在网络的末端用亚像素插值实现超分辨率。Zhou 等人(2018a)直接运用残差网络学习

低分辨率的阵列图像到对应高分辨率图像的映射关系。Xu 等人(2020)提出预去马赛克网络(pre-demosaicing network, PDNet)和残差密集压缩网络(residual-dense squeeze-and-excitation networks, RD-SEN)的级联网络,首先将 RI 插值得到的 RGB 图作为替代目标,与 PDNet 得到的结果进行残差,得到中间的去马赛克结果,然后利用 RDSen 学习中间去马赛克结果和参考图像的映射关系,并进行感知优化。

### 3 实验

#### 3.1 数据集和评价指标

在图像去马赛克任务中,常用的训练数据集有 WED(Waterloo exploration database)数据库、ImageNet、DIV2K(DIVERse 2K resolution high quality images)和 Flickr2K。WED 数据库中包含 4 744 幅不同场景的高质量彩色图像。ImageNet 是目前世界上图像识别最大的数据库,包含 1 000 种类别的对象组成的 13 万幅图像,涉及大量不同场景和结构。DIV2K 包含 1 000 幅 2 K 分辨率的高清图。Flickr2K 包含 2 650 幅 2 K 的高清图。常用的测试集主要有 Kodak 和 McMaster(IMAX)。Kodak 数据集中包含 24 幅  $765 \times 512$  像素的图像,以建筑物、天空和大海为视角。McMaster 包含 18 幅从高分辨率图像中裁剪出的大小为  $500 \times 500$  像素的图像,内容主要以室内场景为主。

使用峰值信噪比(PSNR)和结构相似度(structural similarity index, SSIM)作为评价指标。PSNR 主要是以像素的形式衡量重建图像与原始图像的差异。SSIM 主要衡量重建图像与原始图像之间的相似程度。PSNR 和 SSIM 的值越大,表示重建质量越好,方法性能越好。

#### 3.2 实验结果与分析

实验在 Kodak 和 McMaster 两个基准数据集上进行,对具有代表性的双线性插值(Hou 和 Andrews, 1978)、RI(Kiku 等, 2013)、ARI(Monno 等, 2015)等传统方法和 CDM-CNN(Tan 等, 2017)、3-stage(Cui 等, 2018)、NTSDCN(Wang 等, 2021b)、U-Net++(Wang 等, 2021a)等典型深度学习方法进行评价。为保证公平性,测试均在同一个设备上,并且基于深度学习的方法均在 CPU 上进行。

#### 3.2.1 客观评价实验结果

表 1 和表 2 分别是 7 种方法在 Kodak 和 McMaster 数据集上的性能比较。表 3 是测试时间的比较。可以看出:1)大部分基于深度学习方法的精度明显比传统方法高。在 Kodak 数据集上,深度学习中精度最高的 CDM-CNN 的 PSNR 比传统方法中精度最高的 ARI 高出 2.72 dB,同时比 ARI 的时间成本低。该结果证明,CDM-CNN 方法充分利用了 RGB 三通道间的相关性以及绿像素的高采样率。2)双线性插值方法的时间成本最小,但重建效果最差。对于 U-Net++ 方法,虽然精度比 CDM-CNN 在两个基准数据集上分别低 0.62 dB 和 1.27 dB,但在 CPU 上的测试时间提高了 90.71% 左右。实验证明,U-Net++ 在精确度和时间成本上能实现较好的平衡。

表 1 不同算法在 Kodak 数据集上的性能对比  
Table 1 Comparison of performance for different algorithms on the Kodak dataset

| 算法                         | PSNR/dB      | SSIM         |
|----------------------------|--------------|--------------|
| 双线性插值(Hou 和 Andrews, 1978) | 29.45        | 0.890        |
| RI(Kiku 等, 2013)           | 37.88        | 0.979        |
| ARI(Monno 等, 2015)         | 39.21        | 0.981        |
| CDM-CNN(Tan 等, 2017)       | <b>41.93</b> | <b>0.988</b> |
| 3-stage(Cui 等, 2018)       | 40.63        | 0.986        |
| NTSDCN(Wang 等, 2021b)      | 37.25        | <b>0.988</b> |
| U-Net++(Wang 等, 2021a)     | 41.31        | 0.987        |

注:加粗字体表示各列最优结果。

表 2 不同算法在 McMaster 数据集上的性能对比  
Table 2 Comparison of performance for different algorithms on the McMaster dataset

| 算法                         | PSNR/dB      | SSIM         |
|----------------------------|--------------|--------------|
| 双线性插值(Hou 和 Andrews, 1978) | 34.26        | 0.962        |
| RI(Kiku 等, 2013)           | 36.49        | 0.962        |
| ARI(Monno 等, 2015)         | 37.46        | 0.966        |
| CDM-CNN(Tan 等, 2017)       | <b>38.98</b> | 0.972        |
| 3-stage(Cui 等, 2018)       | 38.85        | <b>0.973</b> |
| NTSDCN(Wang 等, 2021b)      | 34.40        | 0.971        |
| U-Net++(Wang 等, 2021a)     | 37.71        | 0.965        |

注:加粗字体表示各列最优结果。

表 3 不同算法在 Kodak 和 McMaster 数据集上测试时间的比较  
Table 3 Comparison of test time for different algorithms on Kodak and McMaster datasets

| 算法                            | Kodak       | McMaster    | 总时间 /s      |
|-------------------------------|-------------|-------------|-------------|
| 双线性插值<br>(Hou 和 Andrews,1978) | 16.3        | 9.69        | 25.99       |
| RI(Kiku 等,2013)               | 37.1        | 20.4        | 57.5        |
| ARI(Monno 等,2015)             | 739.1       | 338.9       | 1 078       |
| CDM-CNN(Tan 等,2017)           | 277.0       | 126.8       | 403.8       |
| 3-stage(Cui 等,2018)           | 675.7       | 310.9       | 986.6       |
| NTSDCN(Wang 等,2021b)          | 440.0       | 210.7       | 650.7       |
| U-Net ++ (Wang 等,2021a)       | <b>27.4</b> | <b>10.1</b> | <b>37.5</b> |

注:加粗字体表示深度学习方法中各列最优结果。

3.2.2 主观评价实验结果

对于图像去马赛克任务来说,最大的挑战是在纹理丰富和色彩过渡区域的重建。图 8 和图 9 分别是在 Kodak 第 1 幅和 McMaster 第 8 幅上不同方法重建效果的局部放大图。可以看出,传统方法极易产生边缘伪影和拉链效应,而深度学习无明显边缘伪影,即使是 PSNR 较低的 NTSDCN(Wang 等,2021b)算法也极少出现伪影,实现了较好的视觉效果。

4 结 语

图像去马赛克是计算机视觉领域中基本的图像处理任务之一,也是图像分类、目标检测等任务的基础。传统图像去马赛克方法中,重建图像多存在颜

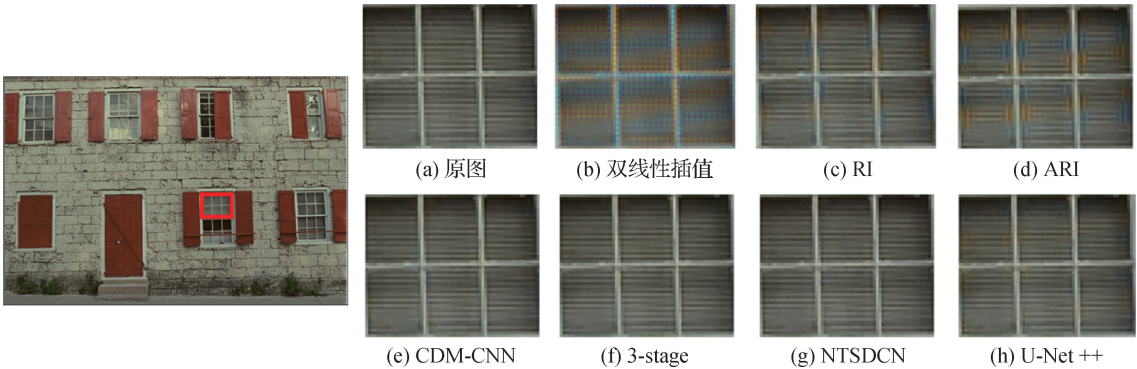


图 8 Kodak 上不同方法重建图像的局部放大图

Fig. 8 The partial enlarged view of the reconstructed image with different methods on Kodak  
( (a) original image; (b) bilinear; (c) RI; (d) ARI; (e) CDM-CNN; (f) 3-stage; (g) NTSDCN; (h) U-Net ++ )

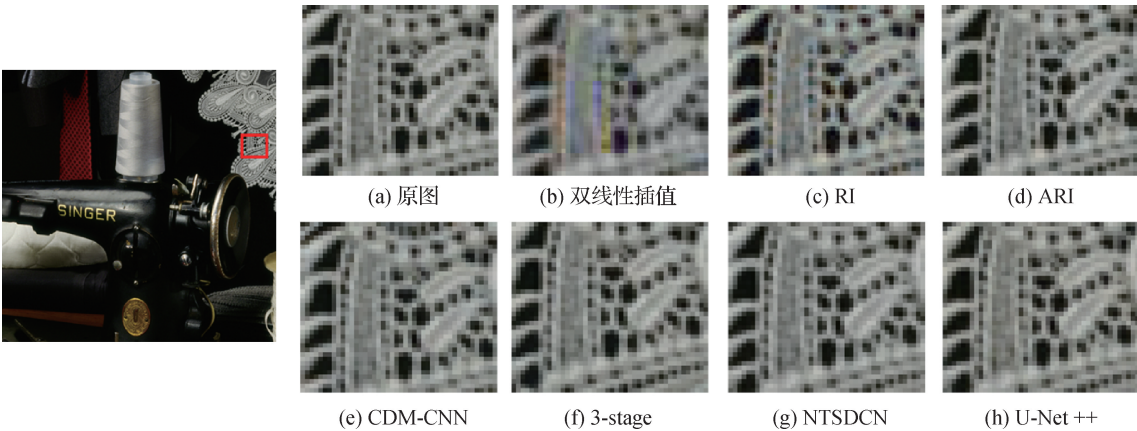


图 9 McMaster 上不同方法重建图像的局部放大图

Fig. 9 The partial enlarged view of the reconstructed image with different methods on McMaster  
( (a) original image; (b) bilinear; (c) RI; (d) ARI; (e) CDM-CNN; (f) 3-stage; (g) NTSDCN; (h) U-Net ++ )



色伪影、拉链效应和棋盘效应等缺陷。随着深度学习在去马赛克任务中的应用,提出了多种重建模型,重建精度得到提升。本文主要对 Bayer 阵列图像的去马赛克方法进行分类总结,阐述了不同方法的基本原理和网络结构,比较了不同方法的重建性能。实验和分析证明,深度学习方法在重建精度方面已超越大部分传统方法,但仍存在时间成本高和应用性不强等问题。

根据已有的研究方法和研究思路,本文对基于深度学习的 Bayer 阵列图像去马赛克技术的待解决问题和研究方向进行展望。1) 部分深度学习方法过度注重网络结构设计,易造成网络过深、不收敛等问题。虽然该思路能够改善图像重建质量,但需存储大量模型参数,计算时间较长。因此,在设计网络结构的同时,应当将计算成本考虑在内,如何平衡计算成本和算法的准确性是图像去马赛克任务的一大挑战;2) 在测试集的图像特征与训练集差异较大时,容易造成网络训练结果不理想,导致模型泛化能力弱,所以进行模型训练时,应选择图像特征丰富的数据集,提高模型的泛化能力;3) 部分研究工作只有系统的客观评价指标,没有系统的主观评价机制,例如 NTSDCN (Wang 等, 2021) 能够获得很好的视觉质量,但是 PSNR 不高。在模型评价时,应结合主客观指标综合评价;4) 多数去马赛克方法仅关注软件层面的改进和提高,针对图像获取硬件的研究较少。围绕软硬件结合的图像去马赛克方法是后续的研究方向之一。

## 参考文献 (References)

- Adams J E Jr. 1995. Interactions between color plane interpolation and other image processing functions in electronic photography//Proceedings Cameras and Systems for Electronic Photography and Scientific Imaging. San Jose, USA; SPIE; 144-151 [DOI: 10.1117/12.204825]
- Bayer B E. 1976. Color imaging array. U. S. , No. 3 971 065
- Cui K, Jin Z and Steinbach E. 2018. Color image demosaicking using a 3-stage convolutional neural network structure//Proceedings of the 25th IEEE International Conference on Image Processing. Athens, Greece; IEEE; 2177-2181 [DOI: 10.1109/ICIP.2018.8451020]
- Din S, Paul A and Ahmad A. 2020. Lightweight deep dense demosaicking and denoising using convolutional neural networks. Multimedia Tools and Applications, 79 (45): 34385-34405 [DOI: 10.1007/s11042-020-08908-4]
- Dong C, Loy C C, He K M and Tang X O. 2014. Learning a deep convolutional network for image super-resolution//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland; Springer; 184-199 [DOI: 10.1007/978-3-319-10593-2\_13]
- Dong W S, Yuan M R, Li X and Shi G M. 2018. Joint demosaicking and denoising with perceptual optimization on a generative adversarial network [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/1802.04723v1.pdf>
- Gharbi M, Chaurasia G, Paris S and Durand F. 2016. Deep joint demosaicking and denoising. ACM Transactions on Graphics, 35 (6): #191 [DOI: 10.1145/2980179.2982399]
- Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y. 2014. Generative adversarial nets [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/1406.2661.pdf>
- Guo Y, Jin Q Y, Facciolo G, Zeng T Y and Morel J M. 2020. Residual learning for effective joint demosaicking-denoising [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/2009.06205.pdf>
- He K M, Sun J and Tang X O. 2013. Guided image filtering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35 (6): 1397-1409 [DOI: 10.1109/TPAMI.2012.213]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA; IEEE; 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hou H and Andrews H. 1978. Cubic splines for image interpolation and digital filtering. IEEE Transactions on Acoustics, Speech, and Signal Processing, 26 (6): 508-517 [DOI: 10.1109/TASSP.1978.1163154]
- Huang G, Liu Z, van der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA; IEEE; 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Huang T, Wu F F, Dong W S, Shi G M and Li X. 2018. Lightweight deep residue learning for joint color image demosaicking and denoising//Proceedings of the 24th International Conference on Pattern Recognition. Beijing, China; IEEE; 127-132 [DOI: 10.1109/ICPR.2018.8546057]
- Ioffe S and Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift//Proceedings of the 32nd International Conference on Machine Learning. Lille, France; JMLR.org; 448-456
- Iriyama T, Sato M, Aomori H and Otake T. 2021. Deep demosaicking considering inter-channel correlation and self-similarity. Nonlinear Theory and Its Applications, 12 (3): 453-463 [DOI: 10.1587/nolta.12.453]
- Jia F, Ma L Y, Yang Y J and Zeng T Y. 2021. Pixel-Attention CNN with color correlation loss for color image denoising. IEEE Signal Processing Letters, 28: 1600-1604 [DOI: 10.1109/LSP.2021.



- 3100263]
- Kang M and Jung M. 2020. Low-dimensional manifold model for demosaicking from a RGBW color filter array. *Signal, Image and Video Processing*, 14(1): 143-150 [DOI: 10.1007/s11760-019-01535-z]
- Kang Q, Fu Y and Huang H. 2019. Deep color image demosaicking with feature pyramid channel attention//Proceedings of 2019 IEEE International Conference on Multimedia and Expo Workshops. Shanghai, China; IEEE: 246-251 [DOI: 10.1109/ICMEW.2019.00-79]
- Kiku D, Monno Y, Tanaka M and Okutomi M. 2013. Residual interpolation for color image demosaicking//Proceedings of 2013 IEEE International Conference on Image Processing. Melbourne, Australia; IEEE: 2304-2308 [DOI: 10.1109/ICIP.2013.6738475]
- Kiku D, Monno Y, Tanaka M and Okutomi M. 2014. Minimized-Laplacian residual interpolation for color image demosaicking//Proceedings Volume 9023, Digital Photography X. Electronic Imaging. San Francisco, USA; SPIE: 197-204 [DOI: 10.1117/12.2038425]
- Kim J, Lee J K and Lee K M. 2016. Accurate image super-resolution using very deep convolutional networks//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA; IEEE: 1646-1654 [DOI: 10.1109/CVPR.2016.182]
- Kim J S, Ko K and Kim C S. 2020. Adaptive lattice-aware image demosaicking using global and local information//Proceedings of 2020 IEEE International Conference on Image Processing. Abu Dhabi, United Arab Emirates; IEEE: 483-487 [DOI: 10.1109/ICIP40778.2020.9190936]
- Klambauer G, Unterthiner T, Mayr A and Hochreiter S. 2017. Self-normalizing neural networks [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/1706.02515.pdf>
- Kokkinos F and Lefkimmiatis S. 2018. Deep image demosaicking using a cascade of convolutional residual denoising networks//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany; Springer: 317-333 [DOI: 10.1007/978-3-030-01264-9\_19]
- Kokkinos F and Lefkimmiatis S. 2019. Iterative joint image demosaicking and denoising using a residual denoising network. *IEEE Transactions on Image Processing*, 28(8): 4177-4188 [DOI: 10.1109/TIP.2019.2905991]
- Lei P C, Liu C, Tang J G and Peng D L. 2020. Hierarchical feature fusion attention network for image super-resolution reconstruction. *Journal of Image and Graphics*, 25(9): 1773-1786 (雷鹏程, 刘丛, 唐坚刚, 彭敦陆. 2020. 分层特征融合注意力网络图像超分辨率重建. *中国图象图形学报*, 25(9): 1773-1786) [DOI: 10.11834/jig.190607]
- Li J S J and Randhawa S. 2007. Colour filter array demosaicking using cubic spline interpolation//Proceedings of 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing. Honolulu, USA; IEEE: 1-865-I-868 [DOI: 10.1109/ICASSP.2007.366045]
- Liu J, Zhang W J, Tang Y T, Tang J and Wu G S. 2020a. Residual feature aggregation network for image super-resolution//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA; IEEE: 2356-2365 [DOI: 10.1109/CVPR42600.2020.00243]
- Liu L, Jia X, Liu J Z and Tian Q. 2020b. Joint demosaicking and denoising with self guidance//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA; IEEE: 2237-2246 [DOI: 10.1109/CVPR42600.2020.00231]
- Longère P, Zhang X M, Delahunt P B and Brainard D H. 2002. Perceptual assessment of demosaicking algorithm performance. *Proceedings of the IEEE*, 90(1): 123-132 [DOI: 10.1109/5.982410]
- Luo J R and Wang J. 2020. Image demosaicking based on generative adversarial network. *Mathematical Problems in Engineering*, 2020: #7367608 [DOI: 10.1155/2020/7367608]
- Luo J R, Wang J and Yue G D. 2021. CFA image demosaicking algorithm based on generative adversarial network. *Computer Engineering*, 47(7): 249-256, 265 (罗静蕊, 王婕, 岳广德. 2021. 基于生成对抗网络的 CFA 图像去马赛克算法. *计算机工程*, 47(7): 249-256, 265) [DOI: 10.19678/j.issn.1000-3428.0058242]
- Mei Y Q, Fan Y C, Zhang Y L, Yu J H, Zhou Y Q, Liu D, Fu Y, Huang T S and Shi H. 2020. Pyramid attention networks for image restoration [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/2004.13824v4.pdf>
- Monno Y, Kiku D, Tanaka M and Okutomi M. 2015. Adaptive residual interpolation for color image demosaicking//Proceedings of 2015 IEEE International Conference on Image Processing. Quebec City, Canada; IEEE: 3861-3865 [DOI: 10.1109/ICIP.2015.7351528]
- Niu Y and Ouyang J H. 2020. Cross-channel correlation preserved three-stream lightweight CNNs for demosaicking [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/1906.09884v2.pdf>
- Park B and Jeong J. 2018. Image demosaicking using densely connected convolutional neural network//Proceedings of the 14th International Conference on Signal-Image Technology and Internet-Based Systems. Las Palmas de Gran Canaria, Spain; IEEE: 304-307 [DOI: 10.1109/SITIS.2018.00053]
- Park B and Jeong J. 2019. Color filter array demosaicking using densely connected residual network. *IEEE Access*, 7: 128076-128085 [DOI: 10.1109/ACCESS.2019.2939578]
- Park Y, Lee S, Jeong B and Yoon J. 2020. Joint demosaicking and denoising based on a variational deep image prior neural network. *Sensors*, 20(10): #2970 [DOI: 10.3390/s20102970]
- Pekkucuksen I and Altunbasak Y. 2010. Gradient based threshold free color filter array interpolation//Proceedings of 2010 IEEE International Conference on Image Processing. Hong Kong, China; IEEE: 137-140 [DOI: 10.1109/ICIP.2010.5654327]
- Prakash V N V S, Prasad K S and Prasad T J. 2017. Deep learning approach for image denoising and image demosaicking. *International Journal of Computer Applications*, 168(9): 18-26 [DOI: 10.

- 5120/ijca2017914500]
- Qian G C, Gu J J, Ren J S, Dong C, Zhao F R and Lin J. 2021. Trinity of pixel enhancement: a joint solution for demosaicking, denoising and super-resolution [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/1905.02538v1.pdf>
- Ramanath R, Snyder W E, Bilbro G L and Sander W A. 2002. Demosaicking methods for Bayer color arrays. *Journal of Electronic Imaging*, 11(3): 306-315 [DOI: 10.1117/1.1484495]
- Rafinazari M and Dubois E. 2014. Demosaicking algorithm for the Fujifilm X-Trans color filter array//Proceedings of 2014 IEEE International Conference on Image Processing. Paris, France: IEEE: 660-663 [DOI: 10.1109/ICIP.2014.7025132]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Proceedings of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4\_28]
- Sharif S M A, Naqvi R A and Biswas M. 2021. Beyond joint demosaicking and denoising: an image processing pipeline for a pixel-bin image sensor//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Nashville, USA: IEEE: 233-242 [DOI: 10.1109/CVPRW53098.2021.00032]
- Shi W Z, Caballero J, Huszár F, Totz J, Aitken A P, Bishop R, Rueckert D and Wang Z H. 2016. Real-Time single image and video super-resolution using an efficient sub-pixel convolutional neural network//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 1874-1883 [DOI: 10.1109/CVPR.2016.207]
- Syu N S, Chen Y S and Chuang Y Y. 2018. Learning deep convolutional networks for demosaicing [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/1802.03769.pdf>
- Szegedy C, Ioffe S, Vanhoucke V and Alemi A A. 2017. Inception-v4, Inception-ResNet and the impact of residual connections on learning//Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, USA: AAAI Press: 4278-4284
- Tan D S, Chen W Y and Hua K L. 2018. DeepDemosaicking: adaptive image demosaicking via multiple deep fully convolutional networks. *IEEE Transactions on Image Processing*, 27(5): 2408-2419 [DOI: 10.1109/TIP.2018.2803341]
- Tan R X, Zhang K, Zuo W and Zhang L. 2017. Color image demosaicking via deep residual learning//Proceedings of 2017 IEEE International Conference on Multimedia and Expo. Hong Kong, China: IEEE: 793-798
- Tan Y F, Chang K, Li H X, Tang Z H and Qin T F. 2020. Lightweight color image demosaicking with multi-core feature extraction//Proceedings of 2020 IEEE International Conference on Visual Communications and Image Processing. Macau, China: IEEE: 136-139 [DOI: 10.1109/VICIP49819.2020.9301841]
- Verma D, Kumar M and Eregala S. 2020. Deep demosaicing using ResNet-Bottleneck architecture//Proceedings of the 4th International Conference on Computer Vision and Image Processing. Jaipur, India: Springer: 170-179 [DOI: 10.1007/978-981-15-4018-9\_16]
- Wang H L, Han Z H, Li L, Qiu S and Jin W Q. 2021. Residual and high-frequency replacement based color reconstruction method for SONY-RGBW array. *Laser and Optoelectronics Progress*, 58(20): #2004001 (王海琳, 韩正昊, 李力, 裘溯, 金伟其. 2021. 基于残差和高频替换的 SONY-RGBW 阵列彩色重构方法. *激光与光电子学进展*, 58(20): #2004001) [DOI: 10.3788/LOP202158.2004001]
- Wang S Y, Zhao M X, Dou R J, Yu S M, Liu L Y and Wu N J. 2021a. A compact high-quality image demosaicking neural network for edge-computing devices. *Sensors*, 21(9): #3265 [DOI: 10.3390/s21093265]
- Wang Y, Yin S Y, Zhu S Y, Ma Z, Xiong R Q and Zeng B. 2021b. NTSDCN: new three-stage deep convolutional image demosaicking network. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(9): 3725-3729 [DOI: 10.1109/tcsvt.2020.3040082]
- Wan S D, Tang S, Xie X Z, Gu J, Huang R, Ma B and Luo L. 2021. Deep convolutional-neural-network-based channel attention for single image dynamic scene blind deblurring. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(8): 2994-3009 [DOI: 10.1109/TCSVT.2020.3035664]
- Wu C Z, Chen X, Ji D and Zhan S. 2018. Image denoising via residual network based on perceptual loss. *Journal of Image and Graphics*, 23(10): 1483-1491 (吴从中, 陈曦, 季栋, 管曙. 2018. 结合深度残差学习和感知损失的图像去噪. *中国图象图形学报*, 23(10): 1483-1491) [DOI: 10.11834/jig.180069]
- Wu D, Zhao H T and Zheng S B. 2020. Motion deblurring method based on DenseNets. *Journal of Image and Graphics*, 25(5): 890-899 (吴迪, 赵洪田, 郑世宝. 2020. 密集连接卷积网络图像去模糊. *中国图象图形学报*, 25(5): 890-899) [DOI: 10.11834/jig.190400]
- Xie S N, Girshick R B, Dollár P, Tu Z W and He K M. 2017. Aggregated residual transformations for deep neural networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5987-5995 [DOI: 10.1109/CVPR.2017.634]
- Xing W Z and Egiazarian K. 2021. End-to-End learning for joint image demosaicing, denoising and super-resolution//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 3506-3515 [DOI: 10.1109/CVPR46437.2021.00351]
- Xu X, Ye Y F and Li X. 2020. Joint demosaicing and super-resolution (JDSR): network design and perceptual optimization. *IEEE Transactions on Computational Imaging*, 6: 968-980 [DOI: 10.1109/TCI.2020.2999819]
- Ye W and Ma K K. 2015. Color image demosaicing using iterative residual interpolation. *IEEE Transactions on Image Processing*, 24(12):

- 5879-5891 [DOI: 10.1109/TIP.2015.2482899]
- Yu J H and Yang X M. 2021. Double branch residual network for demosaicing. *Laser and Optoelectronics Progress*, 58(2): #0210012 (余继辉, 杨晓敏. 2021. 双分支残差去马赛克网络. *激光与光电子学进展*, 58(2): #0210012) [DOI: 10.3788/LOP202158.0210012]
- Yu J H, Yang X M and Yan B Y. 2020. Color image demosaicking based on deep convolution neural network. *Application Research of Computers*, 37(S1): 343-345 (余继辉, 杨晓敏, 严斌宇. 2020. 基于深度学习的彩色图像去马赛克. *计算机应用研究*, 37(S1): 343-345)
- Zhang J K, An C and Nguyen T. 2018. Deep joint demosaicing and super resolution on high resolution Bayer sensor data//*Proceedings of 2018 IEEE Global Conference on Signal and Information Processing*. Anaheim, USA; IEEE: 619-623 [DOI: 10.1109/GlobalSIP.2018.8646321]
- Zhang Y L, Li K P, Li K, Zhong B N and Fu Y. 2019. Residual non-local attention networks for image restoration. [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/1903.10082v1.pdf>
- Zhao Y Z, Po L M, Zhang T T, Liao Z B, Shi X, Zhang Y J, Ou W F, Xian P F, Xiong J J, Zhou C and Yu W Y. 2019. Saliency map-aided generative adversarial network for RAW to RGB mapping//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop*. Seoul, Korea (South); IEEE: 3449-3457 [DOI: 10.1109/ICCVW.2019.00428]
- Zhou R F, Achanta R and Süsstrunk S. 2018a. Deep residual network for joint demosaicing and super-resolution [EB/OL]. [2021-10-29]. <https://arxiv.org/pdf/1802.06573.pdf>
- Zhou Z W, Siddiquee M R, Tajbakhsh N and Liang J M. 2018b. UNet++: a nested U-Net architecture for medical image segmentation//*Proceedings of the 4th International Workshop, DLMIA 2018, and the 8th International Workshop, ML-CDS 2018*. Granada, Spain; Springer: 3-11 [DOI: 10.1007/978-3-030-00889-5\_1]

## 作者简介



魏凌云, 1997年生, 女, 硕士研究生, 主要研究方向为计算机视觉和深度学习。

E-mail: weilingyun1222@163.com



孙帮勇, 通信作者, 男, 教授, 主要研究方向为多光谱快速成像技术和图像处理。

E-mail: sunbangyong@xaut.edu.cn