

苏州大学

本科毕业设计(论文)

学院(部)	数学科学学院		
题 目	基于卷积神经网络的水下退化图像增强方法研究		
年 级	19 级	专业	信息与计算科学
班 级	数19信息计算	学号	1948401093
姓 名	张子荷		
指导老师	陈富军	职称	副教授
论文提交日期	2023 年 5 月		

基于卷积神经网络的水下退化图像增强方法研究

摘要

随着计算机视觉与图像增强技术的发展，越来越多的方法被应用于水下图像处理。水下图像不同于在空气介质中拍摄的图像，光线易受到水介质影响发生衍射、散射、吸收、衰减，图像具有颜色失真、细节模糊、对比度和饱和度较低、易产生色偏等特点。

为了使由水下探测设备拍摄回来的图像可以进行下一步处理分析，本文提出 UWCNN-plus 算法，将传统图像增强方法与深度学习理论相结合，使用合成的水下图像数据集训练卷积神经网络模型，并进行了两步后处理，先是用动态阈值白平衡算法进行色彩校正，再用单尺度 Retinex 算法进行去模糊和细节锐化。

本文对水下图像增强前后的图像质量进行了主观定性分析，并选择客观评价指标信息熵，UCIQE，UIQM 进行了客观定量分析。结果表明增强后的水下图像具有更丰富的细节，更强的对比度，更好的色彩表现力，更能还原真实的水下图像。

关键字：水下图像增强；卷积神经网络；色彩校正；去模糊；水下图像质量评估

Research on underwater degraded image enhancement method based on convolutional neural network

Abstract

With the development of computer vision and image enhancement technology, more and more methods are applied to underwater image processing. Underwater images are different from images taken in air media. Light is easily affected by water medium for diffraction, scattering, absorption, attenuation. The image has the characteristics of color distortion, blurred details, low contrast and saturation, and easy to produce color cast.

In order to make the images taken by the underwater exploration equipment can be processed and analyzed in the next step, this paper proposes the UWCNN-plus algorithm, which combines the traditional image enhancement method with deep learning theory. We use the synthetic underwater image dataset to train the convolutional neural network model, and perform two-step post-processing. We first use the dynamic threshold white balance algorithm for color correction. The single-scale Retinex algorithm is then used for deblurring and sharpening of details.

In this paper, the image quality before and after underwater image enhancement is subjectively and qualitatively analyzed, and the objective evaluation index information entropy is selected, and the objective quantitative analysis of UCIQE and UIQM is carried out. The results show that the enhanced underwater image has richer details, stronger contrast, better color expression, and can better restore the real underwater image.

Keywords: underwater image enhancement; convolutional neural networks; color correction; deblurring; Underwater image quality assessment

目录

第 1 章 绪论	- 1 -
第 1.1 节 研究背景及意义	- 1 -
第 1.2 节 国内外研究现状	- 1 -
第 2 章 相关理论基础	- 6 -
第 2.1 节 水下图像退化原理	- 6 -
第 2.2 节 水下图像质量评价标准	- 7 -
第 2.2.1 节 主观分析	- 7 -
第 2.2.2 节 客观分析	- 7 -
第 2.3 节 卷积神经网络	- 8 -
第 2.4 节 激活函数	- 8 -
第 3 章 图像处理方法介绍	- 11 -
第 3.1 节 基于卷积神经网络的图像增强方法	- 11 -
第 3.2 节 图像色彩校正	- 12 -
第 3.3 节 图像去模糊	- 13 -
第 4 章 实验设计与结果分析	- 14 -
第 4.1 节 实验配置	- 14 -
第 4.2 节 实验数据集与数据读取	- 14 -
第 4.3 节 实验设计	- 16 -
第 4.3.1 节 基于卷积神经网络的水下退化图像增强实验	- 16 -
第 4.3.2 节 色彩校正实验	- 19 -
第 4.3.3 节 去模糊实验	- 22 -
第 4.4 节 实验结果与评估	- 24 -
第 5 章 总结与不足	- 27 -
第 5.1 节 总结	- 27 -
第 5.2 节 不足	- 27 -
参考文献	- 29 -
致谢	- 31 -

第 1 章 绪论

第 1.1 节 研究背景及意义

海洋蕴含着丰富的资源，但绝大多数领域还未被人类探索发掘。随着人口数量的增长，陆地资源紧缺，对海洋的开发被提上日程。获取海底图像信息是开发利用海洋资源的前提和基础。海洋资源是我国的重要资源，海洋资源的开发是 21 世纪的重要命题。随着计算机视觉与图像增强技术的发展，越来越多的方法被应用于水下图像处理。由于水下图像不同于在空气介质中拍摄的图像，光线易受到水介质影响发生衍射、散射、吸收、衰减，会导致图像颜色失真、细节模糊、对比度和饱和度较低、易产生色偏。自主水下机器人（AUV）是一款用于水下探测和识别、水下作业的智能机器人。水下机器人从深海拍摄得来的图像往往具有颜色偏差、对比度低、失真模糊等特征。因此，水下退化图像增强技术是基于视觉的水下机器人的重要组成部分。

深度学习逐步兴起并被广泛应用于多个领域，水下图像增强领域也在尝试着用深度学习的方法解决问题。本文旨在提出一种基于卷积神经网络的水下退化图像增强方法，应用于水下探测作为预处理步骤，使水下机器人采集来的图片能够进行进一步加工分析。

第 1.2 节 国内外研究现状

随着计算机视觉与图像增强技术的发展，对水下图像处理方法的研究也越来越深入。由于水下图像不同于在空气介质中拍摄的图像，光线易受到水介质影响发生吸收、衰减、衍射和散射，具有颜色失真、细节模糊、对比度较低、易产生色偏等特点。为了使由水下探测设备拍摄回来的图像可以进行下一步处理分析，人们从基于非成像模型和基于成像模型（IFM）两个角度在水下图像增强和复原方面做出了许多努力。本节回顾了近二十年中水下图像增强和复原的算法和技术。

1 基于非成像模型的水下图像增强方法

不使用成像模型的水下图像增强方法主要通过调整水下图像的像素来提高图像对比度，增强视觉效果，这种方法不需要考虑光在水中传播的光学特性和成像原理。

在水下图像的早期研究中，人们将以往用于图像增强和去雾的算法和技术直接应用在水下图像处理上，后来又出现了许多针对水下图像颜色失真、图像模糊的特点的方法，主要为以下 4 种。

1.1 空域法

通过观察图像的直方图，我们可以发现相较于自然图像，水下图像像素的分布相对集中。因此，人们尝试通过扩大直方图的动态范围来增强水下图像的视觉效果。2013 年，Hitam^[1]等人改进了 CLAHE 算法，提出了混合对比度有限自适应直方图均衡（Mix-CLAHE）以增强水下图像的视觉效果。该算法应用于 RGB 和 HSV 颜色模型，生成两个图像，并通过欧几里得范数相结合。结果表明，所得图像的对比度增强，噪声和伪影减少，图像质量显著提高。

1.2 转换域法

通过将原图像从空间域转换至频域来增强图像的方法就是转换域法。在频域中，高频分量对应像素变化大的边缘区域，低频分量对应像素变化小的背景区域。图像变换至频域后，通过放大高频分量，减小低频分量可以显著提高水下图像的视觉效果。Agaian^[2]等人在论文中提出通过傅里叶变换进行图像增强。然而，水下图像不同于普通图像，图像主体的边缘区域的高频分量和背景区域的低频分量差异较小。针对水下图像的特点，国内外学者尝试使用同态滤波器、小波变换来处理水下图像。滤波器通常用于图像去模糊。Bazeille^[3]等人在论文中将图像从 RGB 空间转换至 YCbCr 空间，通过同态滤波器消除非均匀光照的影响，然后对图像进行小波去噪，再通过异性滤波器增强图像边缘，最后将图像从 YCbCr 空间转换回 RGB 空间得到增强后的图像。

1.3 色彩校正

针对水下图像的颜色失真问题，通过校正 RGB 空间的色彩来进行图像增强。2007 年，Iqbal^[4]等人提出了一种基于集成颜色模型的水下图像增强方法，该算法先在 RGB 空间中拉伸对比度来平衡图像色彩，再在 HSI 空间中拉伸饱和度和亮度来还原图像的真实颜色。图 1 显示了运用 Iqbal 等人的算法前后的图像对比。

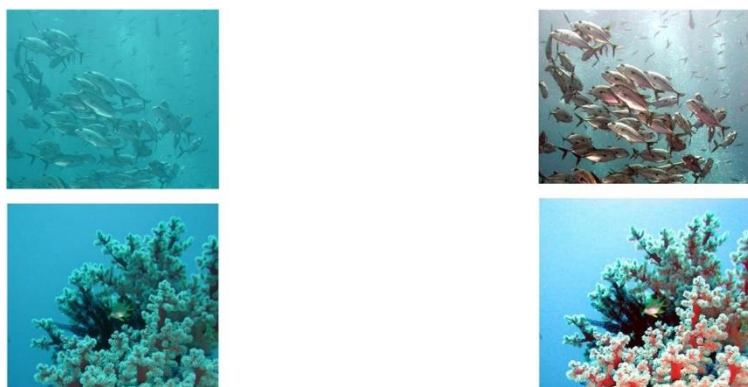


图 1：原始图像和使用 Iqbal 等人方法增强后的图像（图源自 Iqbal）

1.4 融合法

融合原理是将不同状态的目标图像的集合进行信息融合，包含如退化图像、色彩校正图像、对比度增强版本图像等。2012 年，Ancuti^[5]等人提出了一种基于多尺度融合的方法，即将图像的对比度增强版本与白平衡版本作为输入，生成无伪影的图像。2015 年，Khan^[6]等人将融合原理与小波变换相结合，先输入原始图像的白平衡、最小、最大版本，再由图像的亮度、色度、对比度定义融合权重。基于融合原理将输入与融合权重融合后进行小波变换，得到最终结果图像。



图 2：原始图像 a 和复原后图像 b（图源自 Khan）

2 基于成像模型的水下图像增强方法

成像模型通常是指 Jaffe^[7]在论文中建立的光学成像模型，该模型将水下图像分解为三个部分的线性叠加，即水下探测设备接收到的光由三个部分组成：直接光 E_d （未

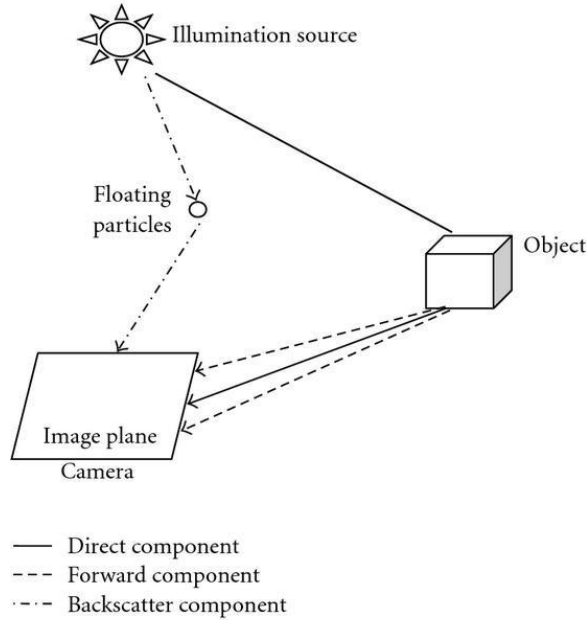


图 3：水下图像的三个组成部分：直接光、前向散射光、后向散射光（图源自 Schettini）

在水中散射的物体反射光）、前向散射光 E_f （以小角度散射的物体反射光）和后向散射光 E_b （并不是直接进入相机，而是通过水中漂浮粒子折射的物体反射光）^[8]。因此得到总照度 E_T 为：

$$E_T = E_d + E_f + E_b$$

光在水中发生吸收、衰减、散射、衍射和水中的悬浮颗粒是水下图像退化的主要原因。针对水下图像的光学特性，人们使用了基于先验的方法来增强水下图像。2011 年，He^[9]等人提出了暗通道先验（DCP）方法，该方法最初被用于图像去雾。由于雾霾图像与水下图像有相似之处，暗通道先验也被广泛用于水下图像增强。红光在水下传播时衰减速度比绿光和蓝光快很多，水下图像的暗通道以红色通道为主。2015 年，Galdran^[10]等人提出了一种基于红通道的算法，是对暗通道法的改进，该算法提高了整体图像的色彩保真度，增强了图像的人工光区域。随着深度学习的发展，研究从通过人工优化模型参数转变为自动训练模型参数。2017 年，Ding^[11]等人在论文中先使用改进白平衡算法提高图片整体质量，再采用卷积神经网络估计水下图像的背景光和

透射率，最后基于成像模型得到了结果图像。

3 结论

直方图均衡、色彩校正、暗通道先验是一些最常用的方法。近几年，引入了多尺度融合技术、卷积神经网络大大减少了计算工作量。

第 2 章 相关理论基础

第 2.1 节 水下图像退化原理

通常，水下采集来的图像会退化是因为光会被水中的三种成分：有色溶解有机物，微型浮游植物和非藻类颗粒吸收和散射。海水对不同波长的可见光具有不同的吸收和散射效应，光的衰减与距离密切相关。在海洋中，波长最长的可见光是最先被吸收的。图 4 展示了不同颜色的可见光在海水中传播的深度，我们可以发现光的吸收与其波长具有一定的相关性。一般来说，红光最先随着距离的增加而消失，其次是橙光、黄光、紫光、绿光和蓝光。因此，我们看到的拍摄到的水下图片往往以蓝色或绿色为主，红色所占的部分较少。光在复杂的水下环境中传播时，水下探测设备接收到的光由三个部分组成：直接光 E_d （未在水中散射的物体反射光）、前向散射光 E_f （以小角度散射的物体反射光）和后向散射光 E_b （并不是直接进入相机，而是通过水中漂浮粒子折射的物体反射光）。因此，为了提高水下采集来的图像的视觉效果和质量，我们需要找到一种能够消除光线散射和依赖于波长的光吸收的影响的方法。

比较同样退化效应较为严重的雾霾天气图像和水下图像，我们可以发现二者有较多的相似之处，因此一些图像去雾理论可以类比地应用到水下退化图像的增强与复原上。

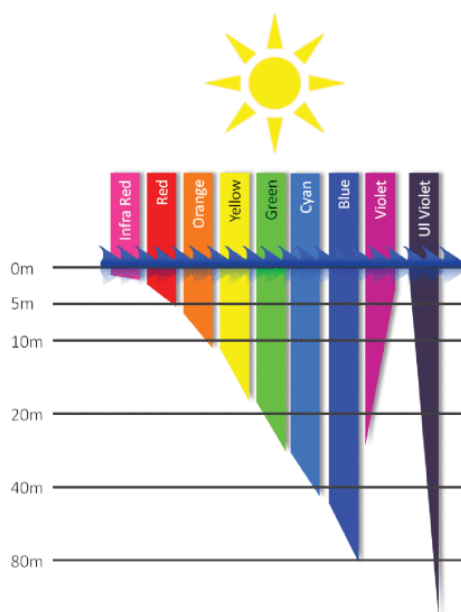


图 4：吸收可见光（图源自 S. -C. Pei and C. -Y. Chen）

第 2.2 节 水下图像质量评价标准

图像质量常常会受到成像设备的光学性能、成像条件、仪器误差、图像处理等因素的影响。图像质量评估通常分为主观定性评估和客观定量评估。在过去的几年中，人们已经提出并分析了许多不同的图片质量评估方法，目的是得到与感知图像质量测量相关的评价指标。

第 2.2.1 节 主观分析

主观定性评估主要依靠人类视觉系统来获得对图像的主观印象。一个合适的主观评估需要通过改变影响图片质量的因素进行重复实验来生成一个数据集，然后由人类观察者评分，统计结果。由于主观定性评估效率低、操作复杂、受观察者主观影响较大，本文仅以不同的水下图像增强和复原方法的代表性结果作为主观分析的基础。

第 2.2.2 节 客观分析

客观定量评估通过建立基于人类视觉系统的数学模型来计算质量指标。只要使用准确可靠的模型，这种方法通常比主观定性评估更为有效。

峰值信噪比和均方误差是目前最广泛使用的客观评价失真图像质量的指标。近几十年中人们已经做出了巨大努力开发新的客观评价图像质量的方法，并通过考虑人类视觉系统特征结合感知质量测量。

信息熵（Entropy）也是一种客观评价增强前后图像质量的指标。熵指的是体系的混乱的程度，对焦良好的图像的信息熵大于没有对焦清晰的图像，因此可以用信息熵作为一种图像清晰度评价标准。信息熵越大，图像越清晰。

水下退化图像增强是为了提高图像的色彩对比度、饱和度、可见度。Yang 和 Sowmya^[12]提出了一种水下彩色图像质量评估标准，也就是 UCIQE。在这个评价标准中，他们量化了图像的不均匀色彩偏差、模糊程度和低对比度，然后以线性方式将这三个分量进行组合。Panetta^[13]等人提出了一种非参考水下图像质量测量标准，称为

UIQM。它包括了三个测量属性：色彩测量、对比度测量和清晰度测量。每个测量属性都是根据人类视觉系统的属性类比得到的。

第 2.3 节 卷积神经网络

卷积神经网络（Convolutional Neural Networks, 简称为 ConvNet 或 CNN）是一种前馈神经网络。它是基于生物学中感受野的概念提出来的。在人类视觉神经系统中，一个神经元的感受野是指视网膜上的特定区域，只有受到该区域的刺激才能够激活这个神经元。ConvNet 是一种端到端的深度学习模型，它拥有自己的特征检测机制。一个训练有素的深层卷积神经网络会以一种分层次的方式自动识别和检测特征，从简单的局部到复杂的整体，如鱼、海藻等。

卷积神经网络一般由卷积层、汇聚层（又叫池化层）和全连接层构成。它的三个结构上的特征为：局部连接、参数共享机制、由多卷积核构成。

卷积层

卷积层能够提取目标图像的一个局部区域的特征，卷积核相当于特征提取器，使用不同个数、不同大小的卷积核都会对特征提取的效果产生影响。

池化层

池化层能够进行特征选择，减少特征的数量，从而可以减少参数的数量，提高效率。

全连接层

全连接层可以作为分类器，全连接层的每一个结点都与上一层的所有结点相连，对前面卷积层和池化层提取到的特征进行综合处理。

第 2.4 节 激活函数

在网络结构的每个神经元中，我们可以将计算后的结果输入到一个非线性函数中再进行输出，这个函数就是激活函数。如果没有激活函数，那么每一层神经网络都是线性函数，由多层网络叠加后仍然是线性函数。激活函数的引入增强了卷积神经网络

的复杂度和表达能力，在卷积神经网络的构建中起到了重要的作用。

Sigmoid 型函数

常用的 Sigmoid 型函数包含 Logistic 函数和 Tanh 函数。

Logistic 函数定义如下：

$$\sigma(x) = \frac{1}{1 + \exp(-x)}$$

Tanh 函数定义如下：

$$\tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$$

Tanh 函数也可以看作是对 Logistic 函数进行放大和平移后，值域变为 $(-1, 1)$ 的 Logistic 函数。它的定义如下：

$$\tanh(x) = 2\sigma(2x) - 1$$

ReLU 函数

ReLU 函数是一种斜坡函数，它的定义如下：

$$\begin{aligned} \text{ReLU}(x) &= \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \\ &= \max(0, x) \end{aligned}$$

ReLU 函数与 Logistic 函数和 Tanh 函数相比，不会出现饱和的情况，也不会造成梯度消失的现象。因为 ReLU 函数中不涉及指数运算，计算效率得到提高，实际的收敛速率与 Logistic 函数和 Tanh 函数相比要快六倍左右。

若我们定义网络和激活函数如下：

$$z = wx + b$$

$$\text{ReLU} = \max(0, z)$$

我们可以得到：

$$\begin{cases} \frac{\partial L}{\partial \text{ReLU}} \frac{\partial \text{ReLU}}{\partial z} \frac{\partial z}{\partial w} = \frac{\partial L}{\partial \text{ReLU}} x, & z > 0 \\ \max(0, z) = 0, & z < 0 \end{cases}$$

当 $x > 0$ 或者 $x < 0$ 时，参数更新的方向与符号相同，故 ReLU 函数和 Sigmoid 型函数一样的缺点，ReLU 函数的输出不以零为中心，参数更新的效率较低。此外，当 $z < 0$ 时，信息在神经网络中无法传递下去。反向传播的过程中，参数梯度恒为零，

无法进行更新，导致了 ReLU 神经元在训练过程中死亡。为了避免上述情况，我们可以使用 ReLU 函数的变形，如带泄露的 ReLU 函数。

带泄露的 ReLU 函数

带泄露的 ReLU 函数即 Leaky ReLU 函数，它的定义如下：

$$\begin{aligned} \text{Leaky ReLU}(x) &= \begin{cases} x, & x > 0 \\ \gamma x, & x \leq 0 \end{cases} \\ &= \max(0, x) + \gamma \min(0, x) \end{aligned}$$

其中 γ 是一个较小的常数，我们可以取为 0.2。当 $\gamma < 1$ 时，带泄露的 ReLU 函数可以写为如下形式：

$$\text{Leaky ReLU}(x) = \max(x, \gamma x)$$

Leaky ReLU 函数将负轴区域设为非零，有效解决了 ReLU 函数易造成“死神经元”的问题，是对 ReLU 函数的重要提升。

MaxOut 函数

MaxOut 函数是一种分段线性函数，它的拟合能力很强，任意凸函数都可以由分段函数逼近，MaxOut 函数就是取 K 个神经节点的最大值，每个节点都是线性函数，它的定义如下：

$$\text{MaxOut}(x) = \max_{k \in [1, K]} (z_k)$$

第 3 章 图像处理方法介绍

第 3.1 节 基于卷积神经网络的图像增强方法

虽然基于深度学习的方法在计算机视觉领域取得了巨大成功,但在基于图像的水下图像增强领域,进展仍然相当有限。水下退化图像与有雾图像在一定程度上有许多相似特征,因此一些图像去雾技术也适用于水下退化图像增强。基于卷积神经网络的图像去雾、细节锐化、图像复原等任务对于水下退化图像增强具有借鉴意义。

DehazeNet 是由 B. Cai^[14]等人提出的一种用于图像去雾的端到端的网络训练模型,这篇文章将传统的图像去雾方法与深度学习相结合,取得了非常好的去雾效果。他们所提出的 DehazeNet 模型由级联卷积层和池化层组成,并且在其中一些层后使用了新的非线性激活函数,这种激活函数叫双边整流线性单元 (BReLU),它是对 ReLU 函数的一种改进和创新。DehazeNet 算法的特征提取不同于传统的卷积神经网络,采用卷积和 MaxOut 单元相结合的结构作为网络的第一层,如下所示:

$$F_1^i(x) = \max_{j \in [1,k]} g^{i,j}(x), g^{i,j} = W_1^{i,j} * I + B_1^{i,j}$$

其中的 W_1 和 B_1 分别代表滤波器和偏置,并进行卷积。

在 DehazeNet 的第二层中用到了用到了多尺度映射,多尺度特征可以提高在不同分辨率下特征提取的鲁棒性,使用相同数量的 3×3 , 5×5 和 7×7 的滤波器,输出如下所示:

$$F_2^i = W_2^{\lceil i/3 \rceil, (i/3)} * F_1 + B_2^{\lceil i/3 \rceil, (i/3)}$$

局部极值 (Max Pooling) 常用于深度神经网络,可以有效抑制噪声,在 DehazeNet 的第三层中使用到了局部极值计算,如下所示:

$$F_3^i(x) = \max_{y \in \Omega(x)} F_2^i(y)$$

为解决 Sigmoid 函数和 ReLU 函数在训练过程中会发生的问题,文中创新地使用了 BReLU 函数,使其在双边受到约束的同时,保证了局部的线性性质。在使用双边整流线性单元的基础上,将 DehazeNet 的第四层定义为:

$$F_4 = \min(t_{\max}, \max(t_{\min}, W_4 * F_3 + B_4))$$

DehazeNet 较好地做到了将传统图像去雾模型与深度学习算法相结合,实现了对

去雾算法的进一步优化，为图像增强提供了解决问题的新思路。本文将在深入研究 DehazeNet 模型后，利用迁移学习构建 UWCNN 网络模型，对水下退化图像进行增强和复原。迁移学习发生在两个不同的领域之间，一个领域已有的知识可以帮助完成新领域上的学习任务。

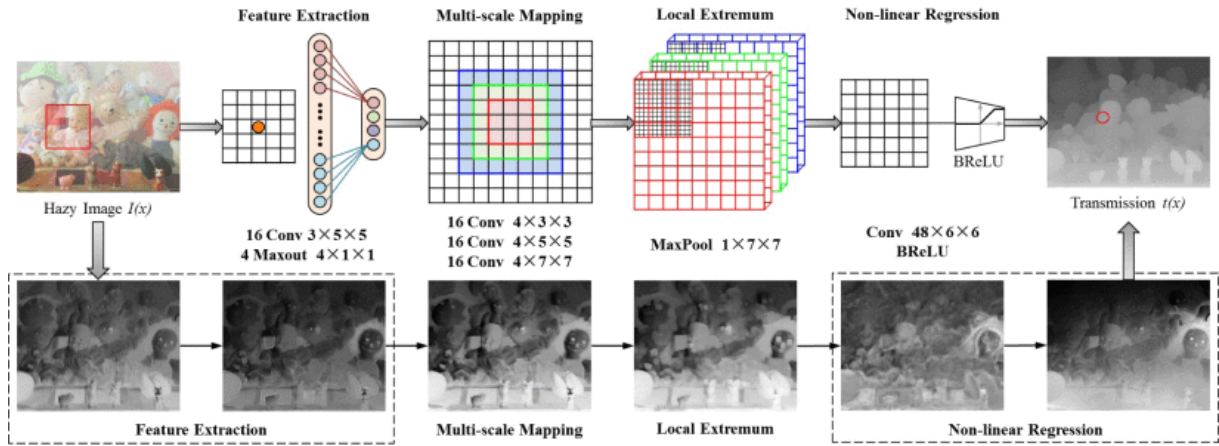


图 5: DehazeNet 网络结构（图源自 B. Cai）

第 3.2 节 图像色彩校正

本文选择白平衡算法对图像进行颜色校正的后处理。白平衡算法能更好地调整图像的颜色畸变，获得细节更清晰、色彩更鲜艳的图像，这种方法也可以应用到水下图像的增强与复原，有助于改善低质量、低照度的水下退化图像，获得更为真实的视觉效果。

白平衡算法是基于如下假设，在单幅图像的 RGB 通道中白色区域对应着最大像素值，暗色区域对应着最小像素值。动态阈值白平衡算法是将图片从 RGB 空间转换到 YCbCr 空间，计算 Cb、Cr 的均值 M_b , M_r 和 Cb、Cr 的均方差。根据阈值的要求提取出近白区域的像素点，将提取出的像素点从亮度值大的点到小的点依次排列，然后以提取出的近白区域中 10% 的亮度值较大的像素点作为参考白点，可得到参考白点的 RGB 三通道的值。根据这些数据计算出参考白点在 RGB 色彩空间的均值和图像亮度最大值，得到 RGB 色彩空间中三通道的增益，通过增益来对图片的 RGB 三通道进行调整，即可得到增强后的图像。

第 3.3 节 图像去模糊

本文选择 Retinex 算法对图像进行去模糊处理。Retinex 是由视网膜（Retina）和皮层（Cortex）构成的，因此 Retinex 理论又称作视网膜皮层理论。该理论的主要观点是，物体能被观察或接收到的图像信息是由两个因素决定的，物体本身所具有的反射性质和物体所处环境的光照强度。原始图像中所有像素点的动态范围大小是由环境光照强度来决定的，而图像的颜色是由物体本身所具有的反射系数决定的。最后形成的图像可用如下公式表示：

$$I(x, y) = L(x, y) \times R(x, y)$$

式中 $I(x, y)$ 代表观察者或照相机接收到的图像信号， $L(x, y)$ 代表环境光的照射分量， $R(x, y)$ 代表带有图像细节信息的目标物体的反射分量。

Retinex 算法将单幅图像分成两个不同的部分：反射图像层和亮度图像层。我们需要去除光照的影响来保留和复原物体的固有属性和状态。等式两边取对数得：

$$\text{Log}(R(x, y)) = \text{Log}(I(x, y)) - \text{Log}(L(x, y))$$

$L(x, y)$ 由 $I(x, y)$ 和一个高斯核的卷积来近似表示：

$$L(x, y) = I(x, y) * G(x, y)$$

上式中 $*$ 代表卷积运算， $G(x, y)$ 代表高斯核。最后，将 $\text{Log}(R(x, y))$ 量化为 0-255 范围的像素值：

$$R(x, y) = (\text{Value} - \text{Min}) / (\text{Max} - \text{Min}) \times 255$$

即可得到去模糊后的图像。

第 4 章 实验设计与结果分析

第 4.1 节 实验配置

本文在 Windows 10 系统平台上使用了 Tensorflow 作为深度学习的框架，并配置有 Intel(R) Core(TM) i5-9300H 主频为 2.4GHz 的处理器和 8GB 的内存。

第 4.2 节 实验数据集与数据读取

本文使用 C. Li^[15]通过模拟真实的水下环境合成的水下图像数据集，对应的真实地面图像使用的是纽约大学数据集 v2，该数据集是来自室内场景的 RGBD 图像，共有 1449 张图像。使用合成的水下图像是为了与真实图像进行对照，如果使用水下退化图像，我们很难获得水下真实参考图像进行卷积神经网络的训练。我们随机选择 960 张图像作为训练集，将水下图像放在目录./NYU_UW/中，真实图像放在目录./NYU_GT/中，剩余图像作为测试集。通过模型合成前后的图像如图 6 所示：



图 6：室内真实图像与水下合成图像

为了提高计算效率，使用标准库函数和 OnenCV2 库函数统一修改图像尺寸为 460×620 像素。

```
width = 620
```

```
height = 460
```

```

img = cv.imread(r"C:\zzh\4x\UWCNN++\test_images\770.png")
img = cv.resize(img, (width, height))
out_file_name = "770out"
save_path = r"C:\zzh\4x\UWCNN++\test_images"
save_path_file = os.path.join(save_path, out_file_name+".png")
cv.imwrite(save_path_file, img)

```

在预处理阶段，需要将图像文件名和路径生成 csv 文件，方便在模型训练阶段对图像进行读取操作。水下图像的文件名为 uw (1)等，真实图像的文件名为 gt (1)等。

```

def createDataCSV (augdir = "./NYU_UW /", gtdir = "./NYU_GT/", outfile="data.csv"):
    df = pd.DataFrame(columns=['AUGFILE', 'GTFILE'])# 列名称
    for i in range (1, 961):
        indexSTR = str(i)
        df = pd.concat([df, pd.DataFrame.from_records([{'AUGFILE': str(augdir+"uw ("
        (+indexSTR+").bmp"), 'GTFILE' : str(gtdir+"gt (" +indexSTR+").bmp")}]])]
    df.to_csv(outfile, index=False)

```

在进行神经网络模型训练前，我们需要先读取图片文件内容，使用 bmp 图像格式解码，转化为 0 到 1 的浮点值。为此我们定义了函数 parse_function 将包含路径的文件名称映射到对应的图像路径上，返回图像数据。

```

# 使用 bmp 格式解码水下图像
filename_image = tf.image.decode_bmp(filename_image_string, channels=3)
filename_image = tf.image.convert_image_dtype(filename_image, tf.float32)
# 使用 bmp 格式解码真实图像
label_image = tf.image.decode_bmp(label_image_string, channels=3)
label_image = tf.image.convert_image_dtype(label_image, tf.float32)

```

先使用 Pandas 库函数读取 csv 文件，再使用 tf.data 建立数据通道，通过 dataset.map 将一个由 Tensor 构成的 Dataset 映射到另一个由 Tensor 构成的 Dataset。这个映射将由文件名构成的 Dataset 映射到了由图像数据构成的 Tensor 上，以下代码用于在训练

前进行大量图像数据的读取。

```
df = pd.read_csv(datafile)
augfiles = list(df["AUGFILE"])
gtfiles = list(df["GTFILE"])
augImages = tf.constant(augfiles)
gtImages = tf.constant(gtfiles)
# 从水下图像和真实图像的切片创建数据集
dataset = tf.data.Dataset.from_tensor_slices((augImages, gtImages))
# 使用长度与数据集大小相等的 buffer size，打乱数据集，这确保了良好的改组
dataset = dataset.shuffle(len(augImages))
# 从图像文件名中解析像素值，批量处理图片，批大小设为 10
dataset = dataset.map(parse_function).batch(10)
```

第 4.3 节 实验设计

在实验中我们使用了 Tensorflow2.0 作为深度学习的框架，并使用 Matlab R2020a 进行图像后处理。

第 4.3.1 节 基于卷积神经网络的水下退化图像增强实验

本文使用受 Tensorflow 支持的 Keras 接口来构建和训练水下图像增强深层卷积神经网络模型（UWCNN）。

```
def __init__(self):
    super(UWCNN, self).__init__()
    # 16 个大小为 3*3*3 的卷积核，strides 为(1, 1)，padding 为 same 填充边缘
    self.conv1 = Conv2D(16, 3, (1, 1), 'same')
    self.relu1 = ReLU()# 激活函数为 ReLU
    self.conv2 = Conv2D(16, 3, (1, 1), 'same')
    self.relu2 = ReLU()
```

```
self.conv3 = Conv2D(16, 3, (1, 1), 'same')
self.relu3 = ReLU()
self.concat1 = Concatenate(axis=3)# 连接前三个卷积层

self.conv4 = Conv2D(16, 3, (1, 1), 'same')
self.relu4 = ReLU()
self.conv5 = Conv2D(16, 3, (1, 1), 'same')
self.relu5 = ReLU()
self.conv6 = Conv2D(16, 3, (1, 1), 'same')
self.relu6 = ReLU()
self.concat2 = Concatenate(axis=3)

self.conv7 = Conv2D(16, 3, (1, 1), 'same')
self.relu7 = ReLU()
self.conv8 = Conv2D(16, 3, (1, 1), 'same')
self.relu8 = ReLU()
self.conv9 = Conv2D(16, 3, (1, 1), 'same')
self.relu9 = ReLU()
self.concat3 = Concatenate(axis=3)
# 3 个大小为 3*3*3 的卷积核
self.conv10 = Conv2D(3, 3, (1, 1), 'same')
self.add1 = Add()
```

读取的图像是 RGB 三通道图像，因此数据的维度为(长, 宽, 通道), 即(620,460,3)。网络结构分为三个相同的模块，每个模块包含三层由 16 个大小为 $3 \times 3 \times 3$ 的卷积核构成的卷积层 Conv，激活函数为 ReLU。在每个模块后有一个 Concat 层，用于连接所有的卷积层，在网络的末端还有一个卷积层，由 3 个大小为 $3 \times 3 \times 3$ 的卷积核构成，所以整个卷积神经网络的深度为 10 层。为了减少图像边界伪影，我们在网络模型中不使用池化层。此外，我们不使用全连接层和批量归一化，这使网络更加高效，易于训练。

卷积为三个通道的所有元素对应相乘后求和，即 27 个乘积的和，滑动步长为每

次横向和纵向移动一个单位，并进行边缘填充，使输出尺寸与输入尺寸匹配。

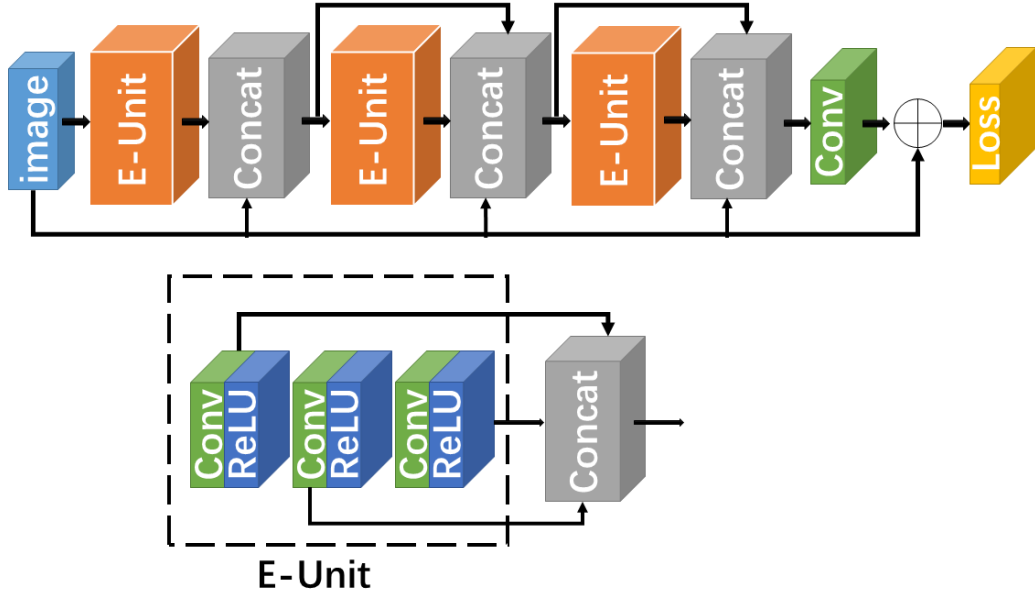


图 7: UWCNN 模型

为了复原水下图像，我们在模型中使用了 L2 损失函数，计算 L2 损失为：

$$\ell_2 = \frac{1}{M} \sum_{i=1}^M |[U(x_i) + \Delta(U(x_i), \theta(x_i))] - I^*(x_i)|^2$$

此外，我们在模型中使用了 SSIM 损失函数，对于图像中每一个像素 x ，在该像素周围一个 13×13 的图像块内计算 SSIM 值：

$$SSIM(x) = \frac{2\mu_{I^*}(x)\mu_I(x) + c_1}{\mu_{I^*}^2(x) + \mu_I^2(x) + c_1} \cdot \frac{2\sigma_{I^*}(x) + c_2}{\sigma_{I^*}^2(x) + \sigma_I^2(x) + c_2}$$

我们将 SSIM 损失函数表示为如下形式：

$$L_{SSIM} = 1 - \frac{1}{M} \sum_{i=1}^M SSIM(x_i)$$

最终的损失函数是 L2 损失函数和 SSIM 损失函数的加和：

$$L = \ell_2 + L_{SSIM}$$

```
def combloss (y_actual, y_predicted):
    # l2 + lssim
    lssim = tf.constant(1, dtype=tf.float32) - tf.reduce_mean(tf.image.ssim(y_actual,
    y_predicted, max_val=1, filter_size=13)) #图像动态范围=1.0
```

```

lmse = MeanSquaredError(reduction=tf.keras.losses.Reduction.SUM_OVER_BATCH_SIZE)(y_actual, y_predicted)
lmse = tf.math.multiply(lmse, 4)
return tf.math.add(lmse, lssim)

```

之后进行模型的训练，并保存模型参数。我们选择 adam 优化器和上述损失函数进行模型优化，迭代次数为 40 次。

```

checkpoint_path = ckptpath
checkpoint_dir = os.path.dirname(checkpoint_path)
# 建立 callback 保存模型参数
cp_callback = tf.keras.callbacks.ModelCheckpoint(filepath=checkpoint_path,
save_weights_only=True, verbose=1)
# 调用模型
model = UWCNN()
model.compile(optimizer=Adam(), loss=combloss)
model.fit(dataset, epochs=40, callbacks=[cp_callback])
# 获取检查点目录
os.listdir(checkpoint_dir)
model.save('save_model/'+type)

```

我们使用的训练数据集是基于室内图像生成的，在经过网络训练后的图像基本去除了蓝绿色调，但图像的动态范围较小，因此我们需要通过色彩校正和去模糊来提高图像的对比度，增强图像边缘和细节。我们将增加了后处理的水下图像增强方法记为 UWCNN-plus。

第 4.3.2 节 色彩校正实验

使用动态阈值白平衡的算法对图像进行色彩校正的后处理，先把图像从 RGB 空间转换到 YCrCb 空间，并取出各分量。设一个参考白点亮度矩阵 RL ，矩阵大小与图

像大小相同。

```
im=imread( img_name);
im2=im;
im1=rgb2ycbcr(im);% 将图片从 RGB 空间转换到 YCrCb 空间
Lu=im1(:,:,1);% 明亮度
Cb=im1(:,:,2);% 色度
Cr=im1(:,:,3);% 色度
[x y z]=size(im);
tst=zeros(x,y);% 亮度矩阵
```

分别计算 Cr、Cb 的平均值 Mr、Mb，根据 Mr、Mb 分别计算 Cr、Cb 的方差 Dr、Db。

```
% 计算 Cb、Cr 的均值 Mb、Mr
Mb=mean(mean(Cb));
Mr=mean(mean(Cr));
% 计算 Cb、Cr 的均方差
Db=sum(sum(Cb-Mb))/(x*y);
Dr=sum(sum(Cr-Mr))/(x*y);
```

判定近白区域，若符合阈值的要求，则可作为参考白点，并把该点(i, j)的亮度值赋给 RL(i, j)，若不符合，则该点 RL(i, j)值为 0。

```
cnt=1;
for i=1:x
    for j=1:y
        b1=Cb(i,j)-(Mb+Db*sign(Mb));
        b2=Cr(i,j)-(1.5*Mr+Dr*sign(Mr));
        if (b1<abs(1.5*Db) & b2<abs(1.5*Dr))% 根据判别式提取出近白区域像素点
            Ciny(cnt)=Lu(i,j);
            tst(i,j)=Lu(i,j);
            cnt=cnt+1;
```



```

        end
    end
end
cnt=cnt-1;

```

选取参考白点中最大的 10% 的亮度值（Y 分量），并选取其中的最小值 Y_{\min} 。
调整亮度矩阵 RL ，若 $RL(i, j) < Y_{\min}$ ，则 $RL(i, j) = 0$ ；否则 $RL(i, j) = 1$ 。

```

%将提取出的像素点从亮度值大的点到小的点依次排列
iy=sort(Ciny,'descend');
nn=round(cnt/10);
%提取出 near-white 区域中 10% 的亮度值较大的像素点做参考白点
Ciny2(1:nn)=iy(1:nn);
mn=min(Ciny2);
for i=1:x
    for j=1:y
        if tst(i,j)<mn
            tst(i,j)=0;
        else
            tst(i,j)=1;
        end
    end
end
end

```

分别把 RGB 三通道的值与亮度矩阵相乘，计算出参考白点的均值和图片亮度的最大值。

```

R=im(:,:,1);
G=im(:,:,2);
B=im(:,:,3);
R=double(R).*tst;
G=double(G).*tst;
B=double(B).*tst;

```

```
% 计算参考白点的 RGB 的均值
Rav=mean(mean(R));
Gav=mean(mean(G));
Bav=mean(mean(B));
Ymax=double(max(max(Lu)))/15;% 计算出图片的亮度的最大值
```

最后，计算出 RGB 三通道的增益，并由增益值调整原图像的 RGB 通道，即可得到色彩校正后的图像。

```
Rgain=Ymax/Rav;
Ggain=Ymax/Gav;
Bgain=Ymax/Bav;
im(:, :, 1)=im(:, :, 1)*Rgain;
im(:, :, 2)=im(:, :, 2)*Ggain;
im(:, :, 3)=im(:, :, 3)*Bgain;
```

第 4.3.3 节 去模糊实验

使用单尺度 Retinex 算法进行图像去模糊，先读取原始图像，分别取出 RGB 三个通道的数值矩阵，转化成 double 型。

```
I = imread('C:\zzh\4x\UWCNN++\results\image.png');
I_r = double(I(:, :, 1));
I_g = double(I(:, :, 2));
I_b = double(I(:, :, 3));
```

然后利用高斯模糊滤波器对各个通道的进行卷积运算处理，将 RGB 图像以及卷积后的图像转化到对数域，再相减。

```
% 将 RGB 图像转化到对数域，拉伸对比度
I_r_log = log(I_r+1);
I_g_log = log(I_g+1);
I_b_log = log(I_b+1);
```

```

% 二维傅里叶变换
Rfft1 = fft2(I_r);
Gfft1 = fft2(I_g);
Bfft1 = fft2(I_b);
% SSR 算法
[m,n] = size(I_r);
sigma = 200;
f = fspecial('gaussian', [m, n], sigma);
efft1 = fft2(double(f));

D_r = ifft2(Rfft1.*efft1);
D_g = ifft2(Gfft1.*efft1);
D_b = ifft2(Bfft1.*efft1);
% 将卷积后的图像转化到对数域
D_r_log = log(D_r + 1);
D_g_log = log(D_g + 1);
D_b_log = log(D_b + 1);
% 将转化到对数域的 RGB 图像和卷积后的图像相减
R = I_r_log - D_r_log;
G = I_g_log - D_g_log;
B = I_b_log - D_b_log;

```

最后，利用指数函数转化回到实数域，对获得的各个通道的图像进行线性拉伸，合并 RGB 通道就可以得到去模糊处理后的图像。

```

R = exp(R);
MIN = min(min(R));
MAX = max(max(R));
R = (R - MIN)/(MAX - MIN);
R = adapthisteq(R);% 改变图像对比度
G = exp(G);
MIN = min(min(G));

```

```

MAX = max(max(G));
G = (G - MIN)/(MAX - MIN);
G = adapthisteq(G);
B = exp(B);
MIN = min(min(B));
MAX = max(max(B));
B = (B - MIN)/(MAX - MIN);
B = adapthisteq(B);
J = cat(3, R, G, B);

```

第 4.4 节 实验结果与评估

将测试集在 UWCNN 上运行前后的图像进行对比，可以发现合成水下图像中的蓝绿色调已经基本去除，图像清晰度提升，无伪影。



图 8：测试集运行前后图像对比

仅仅在测试集上运行模型不足以验证模型的泛化性能，我们又选取了来自真实世界的水下退化图像，进行进一步分析检验。可以看出，该模型的泛化性能较好，在测试集上表现良好，在真实水下图像上的表现仍然很出色。

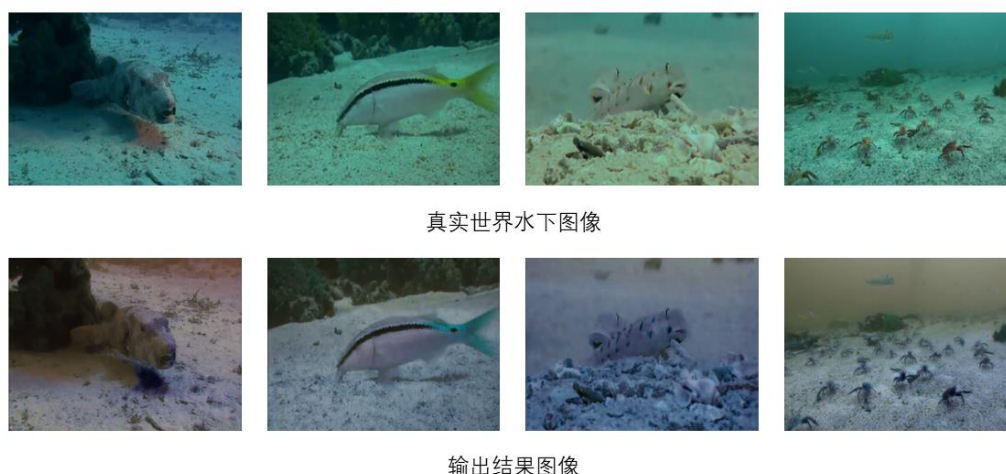


图 9：真实世界水下图像增强前后对比

经过后处理的图像进一步提升了图像的对比度，展现了图像更多的细节，后处理前后图像对比如图 10 所示：



图 10：后处理前后图像对比

我们将本文算法与 HE，CLAHE，defog（暗通道先验去雾）算法进行对比分析来进一步验证它的优越性能。



图 11：算法对比

为了评估水下退化图像增强前后的图像质量，我们进行了主观定性和客观定量的图像评估，并通过人类观察员打分来量化主观评价的结果。给出如上所示的五组图像，我们要求人类观察员从图像清晰度、对比度、真实度、色彩饱和度分别进行打分，权

重相同为 0.25，打分范围为 1 到 10 分，1 分为质量最差，10 分为质量最好，最后得出总分。由所有观察员的给分取平均得到不同图像处理结果的打分。主观评估的得分如下表所示：

	raw	HE	CLAHE	defog	UWCNN-plus
得分	6.43	7.90	8.12	6.89	8.74

其中，本文所用的算法得分排名最高，这表明在经过我们训练好的网络模型和后处理得到的增强后水下图像具有更丰富的细节，更强的对比度，更好的色彩表现力，更能还原真实的水下图像。

图像增强效果的主观分析具有一定局限性，我们选择的客观评价指标为信息熵，UCIQE，UIQM 来进行客观定量分析。

	raw	HE	CLAHE	defog	UWCNN-plus
Entropy	6.808	7.421	7.075	7.136	7.781
UCIQE	0.302	0.337	0.470	0.370	0.413
UIQM	0.236	0.598	0.814	0.495	0.873

从主观分析和客观分析结果我们可以看出，本文的算法具有较好的性能。

第 5 章 总结与不足

第 5.1 节 总结

本文提出 UWCNN-plus 模型，将传统图像增强方法与深度学习理论相结合，对后处理步骤进行创新，先是用动态阈值白平衡算法进行颜色校正，再用单尺度 Retinex 算法进行去模糊和细节锐化。本文提出的图像增强方法基于水下成像模型，充分考虑了水下图像退化的原理，在学习了已有的图像增强方法的基础上，改进后将其用于水下退化图像的增强与复原。

UWCNN 网络模型为一个 10 层神经网络模型，主要由三个相同模块单元构成，每个模块单元由三个卷积层构成，激活函数为 ReLU 函数，模块单元间通过 Concat 层连接，网络末端为一个卷积层进行最终输出。

使用训练好的 UWCNN 模型，经过后处理步骤即可增强和复原水下退化图像。通过主、客观分析结果表明，本文提出的算法明显优于 HE、CLAHE、defog（暗通道先验去雾）算法，具有更丰富的细节，更强的对比度，更好的色彩表现力，更能还原真实的水下图像。

第 5.2 节 不足

卷积神经网络具有局限性，它并不能很好地处理视点变化的其他效果，如旋转和缩放。ConvNet 具有平移不变性，也就是指如果稍微变动同一个物体的位置或者朝向，可能会无法激活那些应该识别该物体的神经元。假如一个神经元是用来识别一只鱼的，那么它的参数就会随着鱼的位置和鱼转动角度的变化而变化。

此外，卷积神经网络无法从目标对象的部分角度来理解图像，它将图像识别为以不同图案排列的像素点，不具有实体及其关系的显式内部表示，会导致一些非常有价值的信息的丢失。同时，网络会忽略图像整体与部分之间的关系。

一个训练好的卷积神经网络的泛化能力是非常重要的评价指标，因此希望以后能构建一个泛化能力更好的网络结构。

由于本文的模型去除了图片中的蓝绿色调，对于水下图像中本来就是蓝绿色的色

块容易产生颜色失真的现象,应该在之后使用更大的数据集或调整模型进行进一步训练。

本文的算法是在 UWCNN 模型的基础上增加了后处理步骤,在这之后尝试构建轻量的单一模型成为了一个可以努力的方向。

参考文献

- [1] Hitam, M.S., Awalludin, E.A., Yussof, W.N.J.H.W., et al.: ‘Mixture contrast limited adaptive histogram equalization for underwater image enhancement’. Proc. Int. Conf. Computer Applications Technology, Sousse, Tunisia, January 2013, pp. 1–5
- [2] S. S. Agaian, K. Panetta and A. M. Grigoryan, "Transform-based image enhancement algorithms with performance measure", IEEE Trans. Image Process., vol. 10, no. 3, pp. 367-382, Mar. 2001.
- [3] Bazeille, S., Quidu, I., Jaulin, L., et al.: ‘Automatic underwater image preprocessing’. Proc. Characterisation du Milieu Marin, France, France, October 2006, pp. 1–11
- [4] Iqbal, K., Abdul Salam, R., Azam, O.: ‘Underwater image enhancement using an integrated colour model’, IAENG Int. J. Comput. Sci., 2007, 34, (2), pp. 1–6
- [5] Ancuti, C.O., Ancuti, C.: ‘Single image dehazing by multi-scale fusion’, IEEE Trans. Image Process., 2013, 22, (8), pp. 3271–3282
- [6] Khan, R.: ‘Underwater image restoration using fusion and wavelet transform strategy’, J. of Comput., 2015, 10, (4), pp. 237–244
- [7] J. S. Jaffe, "Computer modeling and the design of optimal underwater imaging systems, " IEEE Journal of Oceanic Engineering, vol. 15, no. 2, pp. 101–111, 1990.
- [8] R. Schettini and S. Corchs, "Underwater image processing: State of the art of restoration and image enhancement methods, " EURASIP J. Adv. Sig. Pr., vol. 2010, pp. 1-14, 2010.
- [9] He, K., Sun, J., Tang, X., et al.: ‘Single image haze removal using dark channel prior’, IEEE Trans. Pattern Anal. Mach. Intell., 2011, 33, (12), pp. 2341–2353
- [10] Galdran, A., Pardo, D., Picón, A., et al.: ‘Automatic red-channel underwater image restoration’, J. Vis. Commun. Image Represent., 2015, 26, pp. 132–145
- [11] X. Ding, Y. Wang, J. Zhang and X. Fu, "Underwater image dehaze using scene depth estimation with adaptive color correction", Proc. OCEANS, pp. 1-5, 2017.
- [12] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric", IEEE Trans. Image Process., vol. 24, no. 12, pp. 6062-6071, Dec. 2015.
- [13] K. Panetta, C. Gao and S. Agaian, "Human-visual-system-inspired underwater image quality measures", IEEE J. Ocean. Eng., vol. 41, no. 3, pp. 541-551, Jul. 2015.
- [14] B. Cai, X. Xu, K. Jia, C. Qing and D. Tao, "DehazeNet: An End-to-End System for

Single Image Haze Removal," in IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5187-5198, Nov. 2016.

- [15]C. Li S. Anwar and F. Porikli "Underwater scene prior inspired deep underwater image and video enhancement" Pattern Recognit. vol. 98 pp. 107038-107049 Feb. 2020.

致谢

作为苏州大学数学科学学院的本科生，我的本科生活即将收尾。

首先，我要感谢苏州大学，给我们提供了良好的学习环境和住宿条件。苏州大学图书馆资源丰富，为我的论文写作提供了许多帮助。

其次，我要感谢帮助过我的老师和同学们。陈富军老师作为我的论文指导老师，悉心指导我的毕业论文写作，关心我们的论文写作进度，耐心回答我们的问题。陈老师是我《计算机网络》、《Python 程序设计》、《计算机原理》三门课的授课老师，也曾为我写过夏令营推荐信，非常感谢陈老师的帮助、教导和鼓励。陈老师温和风趣，平易近人，是大家心中的好老师。我也要感谢数学科学学院的其他所有老师，让我体会到了数学的奥妙。我还要感谢在数学科学学院认识的同学朋友，在你们身上我看到了许多宝贵的品质，让我也成长为一个更好的人。

最后，我要感谢我的父母，他们在背后给予我力量，父母在精神和物质上的支持让我可以安心进行论文写作。

回顾在苏州大学学习的四年，有很多的收获，不论是知识，还是情谊，都会成为我一生的财富。几个月后，我将成为一名硕士研究生，希望今后能不负自己，不负他人，不负时光。