

基于深度学习的水体色度分类算法研究^①

陈 勇^② 陈 键 陈 焱 裴 植 易文超 王 成 张文珠^③

(浙江工业大学机械工程学院 杭州 310023)

摘 要 运用深度学习技术进行非接触、快速水体色度检测与分类,采用无人机采集水体图像,运用色度仪对标定的图像完成分类,建立数据集。采用图像归一化处理减少环境因素对分类结果的影响,设计多特征的分步边缘检测算法,检测水域图像边缘,剔除无关像素。对 VGG 16、GoogleNet-V3 和 ResNet 18 卷积神经网络进行水体色度分类模型构建与训练,后筛选 Inception 结构和残差结构为基本构建单元,设计专门用于水体色度分类的 WCNet 15 与 WCNet 21 神经网络模型。在训练集上训练参数并利用验证集完成对 2 个模型的准确率的比较,筛选准确率高的 WCNet 21 模型作为最终水体色度分类模型。WCNet 21 模型的最优准确率可达 97.8%,满足水体色度分类需求,可应用到具体的水体色度分类工作当中。

关键词 水质检测;色度分类;深度学习;边缘检测;神经网络

0 引 言

色度作为水体颜色强度的一种表征物理量,常被人们用于评价水体水质的好坏^[1-3]。之前国内外对水体色度的测定方法主要包括稀释倍数法^[4]和铂钴比色法(即目视比色法)^[5]。其中稀释倍数法是将测量水样用蒸馏水或者超纯水稀释到刚好看不到颜色,通过稀释倍数来表示水样色度。铂钴比色法是将测量溶液与预先配制好的铂-钴标准液进行肉眼比色,确定溶液色度值。这 2 种方法都是通过人眼进行目测,存在主观性强、误差大等问题^[6]。随后国内外开始采用分光光度法进行检测,利用特征波长的吸收光谱吸收光峰面积作为色度的衡量标准进行水体色度的检测^[7],其在准确度上有显著提升,且用时短、效率高,能够避免视觉差异对结果产生的影响^[8]。但是它同样需要进行水样的采集与标准色度溶液的配制工作,需要人为参与,程序较为

复杂^[9]。后来光纤传感器的发展使得国内外开始采用光纤传感法进行水体色度测定^[10],该方法利用处理过后的光纤作为传光传感原件,在光的接收端使用光谱仪或光电转换器作为光信息接收端,进而进行水体色度的测试^[11-12]。此方法优点在于不需要准直系统,缺点为传感器稳定性较差^[13]。近些年来,图像传感法成为了一种较为创新的水体色度测定方法,此方法将拍摄获取的图片传输到计算机,再利用图像处理技术获取水体色度信息。如李文等人^[14]提出了一种基于远程图像色度检测点源的方法,此方法精度高,但技术要求也高。

为了进一步拓展图像传感法,本文进行了基于深度学习的水体色度分类方案设计。本方案主要分为两部分,方案的前部分为输入图像的图像预处理模块,包括图像归一化等。方案的后部分为新的水体色度分类模型的搭建。搭建过程中,先对多种卷积神经网络进行水体色度分类模型构建与训练,筛选出能用于水体色度分类的有效结构,后借助有效

① 国家自然科学基金(71871203,51305400)和浙江省科技厅公益项目(LGG22G010002)资助。

② 男,1973 年生,博士,教授;研究方向:智能算法;E-mail: cy@zjut.edu.cn。

③ 通信作者, E-mail: wzzhang@zjut.edu.cn。

(收稿日期:2021-06-28)

结构设计出新的水体色度分类模型(WCNet15 与 WCNet21 模型)。最后进行多个模型的水体色度识别效果比较,对水体分类方案进行分析与总结。

1 色度数据库搭建

采用大疆-MAGIC-PRO 型号标准无人机在目标水域上空 30 ~ 40 m 高度范围内对目标水域进行垂直图像拍摄来采集数据,同时进行相应水域的水样采集工作。使用 BSD-500 微型铂-钴比色仪对获得水样进行铂钴比色法水体色度检测。实验中按照铂-钴标准比色法 GBT 5750.4—2006 1.1 方法进行铂-钴比色法标准液配置,后进行比色仪的标定操作。图 1 为标准无人机拍摄到的部分水域图像数据集。



图 1 水域图像数据集

为了使实验中的水样色度值更加可靠准确,进行如下处理:每份水样的水体色度检测 5 遍,再进行取均值处理,后使用 3 倍标准差法来剔除检测中可能产生的异常数据^[15],进而减少异常数据带来的实验误差。式(1)与(2)为最终色度均值的相关计算公式,其中 $\overline{ch}$ 表示 5 次检测后的色度平均值, θ 表示 5 次检测后的色度标准差, $\overline{ch}_{\text{sati}}$ 表示采用 3 倍标准差法后得到的最终色度均值。

$$\theta = \frac{\sqrt{\sum_{i=1}^5 (ch_i - \overline{ch})^2}}{5} \tag{1}$$

$$\overline{ch}_{\text{sati}} = \frac{\sum_{i=1}^n ch_i}{n} \quad (ch_i \in [\overline{ch} - 3\theta, \overline{ch} + 3\theta]) \tag{2}$$

整个采样过程共采集到 128 份水样,为了解决水样色度数据值分散以及水样色度数据有限的问题,

对水样色度数据进行编号分类,以 50 铂钴色度单位(platium-cobalt chromaticity unit, PCU)为单位进行色度数据分组,其中对未出现的数据以及数据量较少的频段不设置组别,剔除掉一些不必要数据之后,最终得到 120 份色度值介于 0 ~ 200 PCU 之间的水样色度数据,并分成 3 组,具体分类见表 1。

表 1 水体色度分类表

编号	组别色度/PCU	数量
1	0 ~ 50	42
2	50 ~ 100	39
3	> 100	39

2 水域图像预处理

2.1 水体图像归一化与灰度化

在使用无人机进行图像采集时,光照等因素会影响水体图像的相关参数,如水体图像的光强、色调及饱和度等。为解决这个问题,对采集到的水体图像进行光强、色调以及饱和度的归一化处理。

以曝光时间为 1/500 s,光圈值为 f/2.2,ISO 速度为 200 的图像作为归一化的模板图片,进行归一化处理。

图 2 为归一化前后的效果对比图,从左至右 3 列分别是模板水样图像、归一化前水样图像以及归一化后水样图像。由前后对比图易知,图像归一化将水样图像的图像风格变换到了标准图像,避免了外部环境因素的干扰。

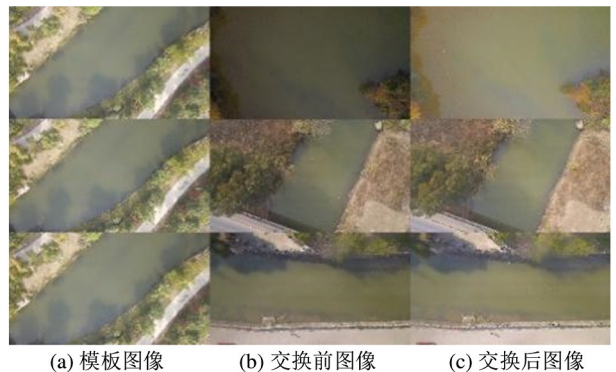


图 2 图像归一化效果图

由于采集到的水体图像均是红绿蓝(red green

blue, RGB)格式的图像,而相较于 RGB 格式的图像,色调饱和度亮度(hue saturation intensity, HSI)格式的图像更适合于计算机进行归一化处理,所以需将采集到的水体图像进行灰度化处理之后再行归一化处理。图3为某一 RGB 格式的水体图像经灰度化后得 HSI 格式水体图像的示意图。

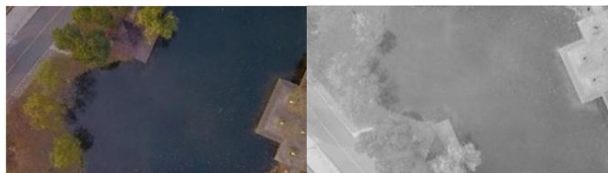


图3 RGB 图像与灰度图像对比

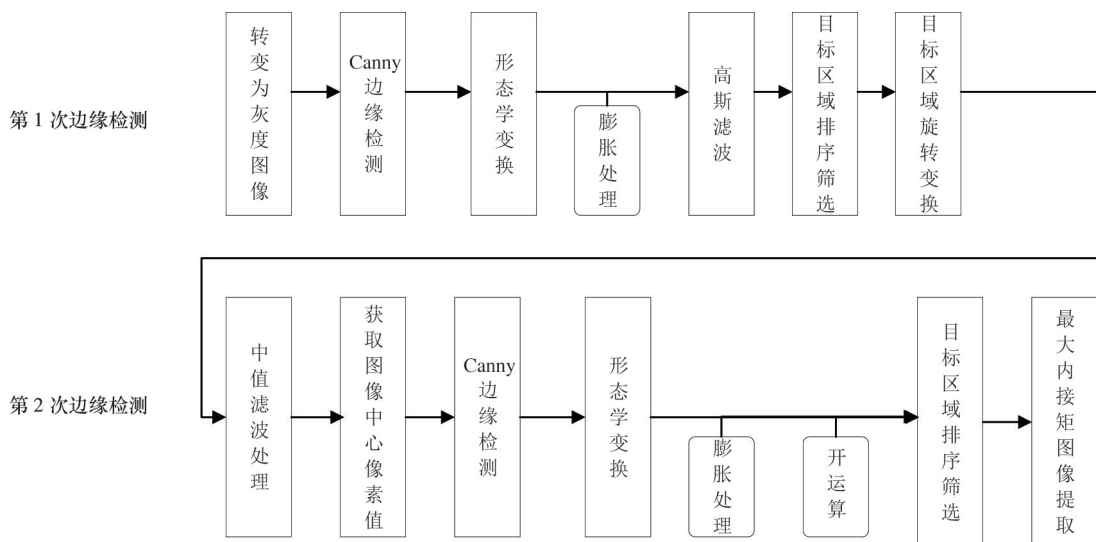


图4 多特征分步边缘检测流程图

树木等干扰像素。由于区分度高,所以在 Canny 算法中可设置较低的梯度下限阈值与上限阈值。对于 Canny 算法而言,其本身包括许多步骤,如高斯平滑、一阶偏导的有限差分计算、非极大值抑制、双阈值算法检测及边缘连接等,Canny 算法的具体步骤如下。

(1) 高斯平滑处理。高斯平滑算法能较好地解决水体图像的噪声问题,先通过图像的二维高斯曲面公式计算出 3×3 图像 $f(a, b)$ 的高斯滤波核 M , 如式(3)所示,其中 (a, b) 表示图像的中心坐标点, σ 为图像 $f(a, b)$ 各坐标点像素值的标准差。后根据式(4),其中 $*$ 表示卷积运算,将高斯滤波核 M 与对应的 3×3 图像 $f(a, b)$ 进行卷积计算。获得处理后的图像 $g(a, b)$, 如此可有效滤去图像当中叠加

2.2 基于多特征的分步边缘检测

在采集到水体图像之后,发现图像中的道路、树木倒影以及水草倒影等因素构成了干扰水体图像识别的干扰像素,又因为这些干扰像素区别较大,难以使用一次性的边缘检测进行剔除。为了解决其对目标水域的干扰,设计一个多特征分步边缘检测操作,操作流程见图4。

对输入的水体图像进行灰度化与归一化处理,后使用 Canny 边缘检测算法^[16]进行第1次边缘检测,此次检测主要针对区分度高的物体,比如道路、

的高频噪声。

$$M = \frac{1}{2\pi\sigma^2} e^{-\frac{a^2+b^2}{2\sigma^2}} \quad (3)$$

$$g(a, b) = M * f(a, b) \quad (4)$$

(2) 采用 sobel 差分算子寻找图像边缘。用 sobel 差分算子计算图像在 x 轴与 y 轴的像素梯度值,像素梯度变化明显的像素点即为图像的边缘,具体运算过程见式(5)~(8)。其中式(5)中的 S_x, S_y 分别为 x 方向和 y 方向上的 sobel 算子, K 表示目标图像中以像素点 $[i, j]$ 为中心的 3×3 灰度矩阵。式(6)中的 G_x 与 G_y 为图像在 $[i, j]$ 像素点处 x 方向与 y 方向的灰度梯度值,其是由 sobel 算子与目标灰度矩阵像素点之间进行卷积计算所获得,具体的 x 方向与 y 方向卷积计算过程见式(6)与(7),其中 $*$ 表示

卷积运算。式(8)中的 G 表示该像素点的总灰度梯度值,其可以用 G_x 与 G_y 的绝对值之和来近似表示。由式(5)~(8)易知,坐标点周围的灰度值发生变化越明显, G 便会越大,反之 G 越小,所以可以用总灰度梯度值 G 作为评判标准来进行相关的边缘检测。

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$K = \begin{bmatrix} a_0 & a_1 & a_2 \\ a_3 & [i, j] & a_5 \\ a_6 & a_7 & a_8 \end{bmatrix} \quad (5)$$

$$G_x = S_x * K \quad (6)$$

$$G_y = S_y * K \quad (7)$$

$$G = \sqrt{G_x^2 + G_y^2} \approx |G_x| + |G_y| \quad (8)$$

(3) 非极大值抑制。对进行过 sobel 算子卷积处理像素点进行局部最大值的寻找,如在一个像素点矩阵 K 中,将 $[i, j]$ 像素点上的像素值与它的 8 邻域像素值进行比较,且 $[i, j]$ 像素点上的总灰度梯度值用 $G_{i,j}$ 表示。若 $G_{i,j}$ 为其中的最大值,则保留其自身灰度值,若不是,便将灰度值设置为 0。

(4) 双阈值检测并连接边缘。设置 2 个梯度阈值 τ_1 和 τ_2 , 其中 τ_2 为 τ_1 的 2~3 倍。使用梯度阈值,对已经经过非极大值抑制处理的图像进行阈值化处理,其中,若像素点的梯度小于阈值,则将像素点的梯度值设为 0,进而可以得到 2 个阈值边缘的图像 I_{τ_1} 和 I_{τ_2} 。相比较之下, I_{τ_2} 比 I_{τ_1} 边缘更准确,因为 I_{τ_2} 是在高阈值得到的。但在其中会出现边缘不连通的问题,所以还需要在 I_{τ_1} 的 8 邻点位置中寻找可连接至轮廓上的边缘,后将该边缘投入 I_{τ_2} 中,用于连通边缘。

经过 Canny 边缘检测之后的水体图像能提取的目标区域有很多,这样便会带来错误提取目标区域的问题,因此在后续步骤中,需对图像进行形态学变换中的膨胀处理与高斯滤波处理。经过上述 2 个处理的结合,可以使得边缘曲线的顺滑度与宽度增加,避免错误提取目标区域的现象发生。具体地,在第 1 次边缘检测时,筛选并提取出无规则的且长度最长的连通目标区域,用于后面进行第 2 次边缘检测。

因为最后设计的神经网络的输入图像为规则的

矩形图像,所以应尽可能保证第 2 次边缘检测获得的内接矩形图像最大。为了使内接矩形图像最大,可对目标区域进行旋转变换。具体操作如下,先寻找出第 1 次边缘检测得到的连通目标区域的凸包节点,以这些节点作为拟合点,后进行直线拟合^[17]。具体拟合过程以及旋转变换的旋转角 $\angle L$ 计算见式(9)~(11)。

$$\rho(r_{c(z)}) = 2 \left(\sqrt{1 + \frac{r_{c(z)}^2}{2}} - 1 \right) \quad (9)$$

$$L: \quad \left(\sum \rho(r_{c(z)}) \right)_{\min} \quad (10)$$

$$\angle L = \frac{\arctan(k) \times 180}{\pi} \quad (11)$$

其中, $r_{c(z)}$ 表示凸包节点 z 到达拟合直线的距离, $\rho(r_{c(z)})$ 表示直线拟合距离函数,当且仅当凸包所有点的直线拟合距离函数之和最小时,获得拟合直线 L , 直线斜率为 k 。

旋转角 $\angle L$ 计算完毕之后便可进行旋转变换,图 5 为某一连通目标区域的旋转示意图。

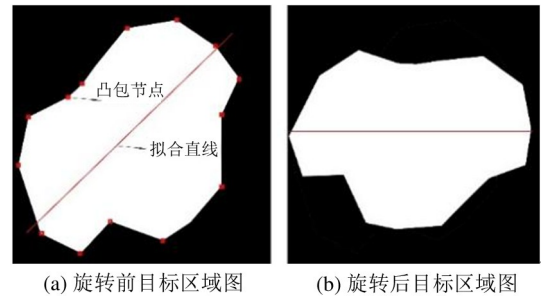
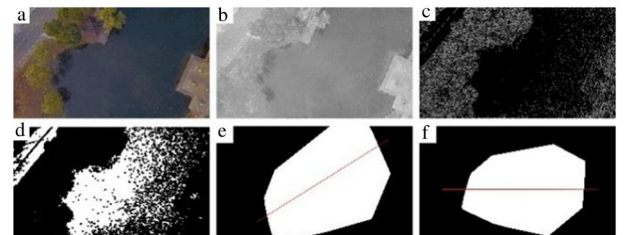


图 5 目标区域旋转变换前后对比图

将第 1 次边缘检测中的 Canny 边缘检测梯度下限阈值设为 10,上限阈值设为 20,后进行实验。图 6 为某一水体图像第 1 次边缘检测的步骤效果图。



(a) 原图; (b) 灰度图像; (c) Canny 边缘检测效果图; (d) 形态学变化效果图; (e) 最大区域提取; (f) 目标区域旋转

图 6 第 1 次边缘检测步骤效果图

在流程方面,第2次边缘检测流程与第1次边缘检测流程基本相同,个别流程环节有差异。第2次边缘检测流程如下。先对第1次边缘检测得到的输入图像进行中值滤波处理,避免目标区域中的颗粒状漂浮物带来像素干扰。再进行 Canny 边缘检测,因为第2次检测的物体(水面倒影、水内水草等)为水体内部干扰像素,再加上水面颜色本身差异不大,所以很有可能在检测过程中剔除掉目标像素。故将 Canny 算法的梯度下限阈值与上限阈值设置的较高一些,其中下限阈值设为 100,上限阈值设为 200,只提取梯度值变化大的像素点作为边缘检测的像素点。后进行形态学处理,在膨胀处理后面增加开运算处理,使得边缘检测的敏感度降低。最后进行最大边缘内的像素提取,以这些像素作为目标像素,再计算最大内接矩形,提取出规则的矩形图像作为最后设计的神经网络的输入图像。图 7 为某一水体图像第2次边缘检测步骤效果图。

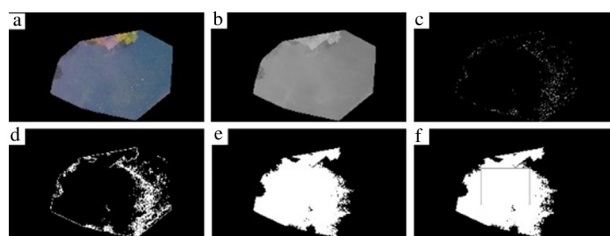


图 7 第 2 次边缘检测步骤效果图

2.3 水体图像数据增强

无人机采集到的且可使用的的水样图像只有 120 张,若只用这些数据作为数据集进行训练,则极易产生过拟合现象。为避免过拟合现象的发生,需对水样图像进行数据增强处理,一张水样图像经过多种空间变换之后可得到更多的水样图像,进而增加数据集当中的数据量。如图 8 为某一水样图像的数据增强处理的示意图,原水样图像进行了水平翻转、纵向翻转、扭曲变换 3 次空间变换,使得图像数据由原先的 1 张变成了现在的 4 张。

对经过数据增强的水体图像进行归一化处理以及多特征分步边缘检测,使得每张水体图像获得一个感兴趣区域(region of interest, ROI)。但对于单张

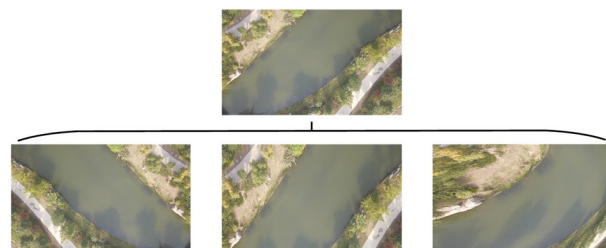


图 8 图像数据增强效果图

水体图像而言,其所包含的目标水域的不同位置会有颜色差异,且其自身图像像素太大,这样会导致神经网络输入数据不可靠,水体色度识别模型不准确。为解决上述问题,在 ROI 区域内部采用随机算子对每张水体图像进行采样工作,式(12)为采样区域的选择,其中 (x_{lu}, y_{lu}) 、 (x_{ru}, y_{ru}) 、 (x_{ld}, y_{ld}) 和 (x_{rd}, y_{rd}) 表示输入水体图像的左上坐标、右上坐标、左下坐标以及右下坐标。 λ_1 与 λ_2 表示在横向与纵向方向上随机产生的左上角横纵坐标的位置, s 表示输入水体图像的像素尺寸, pic_{row} 与 pic_{col} 表示每张水体图像的行数与列数。采样过后,1 张原水体图像将会变成 8 张新的水体图像。

$$\begin{bmatrix} (x_{lu}, y_{lu}) & (x_{ru}, y_{ru}) \\ (x_{ld}, y_{ld}) & (x_{rd}, y_{rd}) \end{bmatrix} = \begin{bmatrix} (\lambda_1, \lambda_2) & (\lambda_1 + s, \lambda_2) \\ (\lambda_1, \lambda_2 + s) & (\lambda_1 + s, \lambda_2 + s) \end{bmatrix} \begin{matrix} \lambda_1 \in (0, pic_{row} - s) \\ \lambda_2 \in (0, pic_{col} - s) \end{matrix} \quad (12)$$

按照上述一系列操作,原有的 120 张水样图像变成了新的 3840 张水样图像,后以训练集:验证集:测试集为 7:1:2 的比例来划分这 3840 张数据集,得到的对应数量比为 2688:384:768,图 9 为部分水体图像输入数据集。

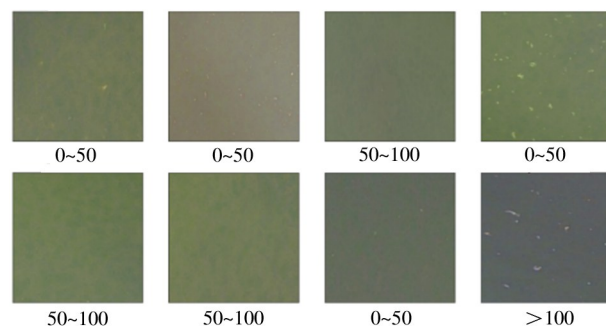


图 9 水体图像输入数据集

3 WCNet 结构选择实验与分析

在水体色度分类模型搭建过程中,先使用 VGG 16、GoogleNet-V3 和 ResNet 18 这 3 类卷积神经网络进行水体色度分类实验,筛选出有利于水体色度分类的结构,后进行分色度类模型的搭建。其中 VGG 16 网络模型的优点在于使用了相同卷积核进行卷积运算,卷积核的串联能很好地扩大感受视野^[18]。GoogleNet 网络模型是由 Szegedy 等人^[19]于 2015 年提出,这是一种全新的分类算法模型,模型中的 Inception 模块被提出。使用 Inception 模块既能控制好参数的数量又能扩展深度学习网络的网络深度^[20]。ResNet 神经网络是由 He 等人^[21]于 2015 年提出的一种网络模型,是一种残差卷积网络,该网络模型不仅提升了自身层数,且面对模型退化问题时,网络模型可以使用一种越级连接结构加以解决。

现进行 3 类卷积神经网络的模型训练,表 2 为 VGG 16、GoogleNet-V3 和 ResNet 18 这 3 类模型的输入数据尺寸。

表 2 各模型输入数据尺寸表

模型类别	输入图像尺寸	通道数
VGG 16	224 × 224	3
GoogleNet-V3	229 × 229	3
ResNet 18	224 × 224	3

为了保证 VGG 16、GoogleNet-V3 和 ResNet 18 三类模型具有可比较性,需对 3 个模型中的相关变量与相关超参数进行设置,采用均值池化层作为 3 个模型的池化层,采用 Relu 函数作为 3 个模型的激活函数,3 个模型的丢弃层(Dropout)均设为 0.5。而对于 GoogleNet-V3 卷积网络模型,还需在每层卷积之后再行批归一化(batch normalization,BN)操作。

三类模型进行训练的优化器均为 Adam 优化器,其中 Adam 方法中的参数 β_1 与 β_2 设为 0.9 与 0.999,数据的批次(Batch)大小设为 8,迭代次数为 5000 次。模型的初始学习率设为 0.01,且每迭代 100 代,学习率变为原来的 1/2。在模型参数初始化操作方面,采用随机初始方式。

采用交叉熵损失函数作为 3 类网络模型的损失函数,同时为避免过拟合现象,在损失函数上增加权重为 0.0006 的 L2 正则项。式(13)为损失函数的具体表达式。

$$J = - \sum_{i=0}^K y_i \log(p_i) + 0.0006 \|w\| \quad (13)$$

其中, J 为损失函数, y_i 为图像标签, p_i 为预测概率, $\|w\|$ 为正则项。

训练完毕后发现 VGG 16 模型的水体色度分类准确率较低,验证集准确率一直维持在 47.91% 左右,同时验证集损失值维持在 1.067 左右。改变初始值后再进行多次实验,发现识别率依旧不高,说明 VGG 16 模型不是很适用于水体色度分类,故后续不再对 VGG 16 模型做进一步的研究。

而 GoogleNet-V3 模型与 ResNet 18 模型的分类型准确率则较高。图 10 为 GoogleNet-V3 模型与 ResNet 18 模型的具体训练结果,其中带有叉号的折线为 GoogleNet-V3 模型的训练结果,带有圆圈的折线为 ResNet 18 模型的训练结果。由于迭代的次数为 5000 次,若不对准确率进行均值处理,则训练结果图会因为准确率波动太大而变得不清楚,所以对每 100 次迭代的准确率进行均值处理,用处理后的准确率进行绘图。

表 3 为 GoogleNet-V3 模型与 ResNet 18 模型在迭代过程当中的最优准确率。由图 10 以及表 3 可知,GoogleNet-V3 模型迭代至 2000 代时准确率趋于稳定,最优准确率为 93.8%;ResNet 18 模型迭代至 1500 代时准确率趋于稳定,最优准确率为 90.6%。就准确率而言,GoogleNet-V3 模型的最优准确率要高于 ResNet 18 模型的最优准确率,且在每一代的迭代过程中,GoogleNet-V3 模型的准确率也基本高于 ResNet 18 模型的准确率。就模型的迭代波动而言,在达到稳定准确率的前提下,GoogleNet-V3 模型的迭代波动要更小。

由 GoogleNet-V3 模型与 ResNet 18 模型训练结果易知,两者在水体色度分类上有着较高的准确率,

表 3 模型最优准确率表

模型类型	GoogleNet-V3	ResNet 18
最优准确率	93.8%	90.6%

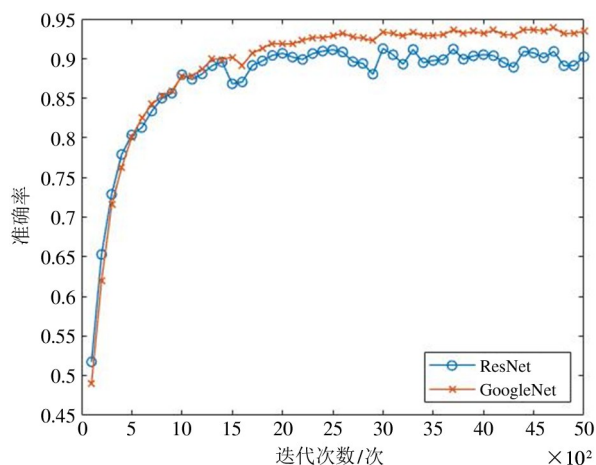


图 10 各模型训练结果图

能很好地进行水体色度的分类。后分析 2 个模型的内部结构,发现残差结构与 Inception-V3 结构可以有效降低参数量以及避免梯度爆炸现象的产生。这 2 个结构有利于水体色度的分类,所以可以用这 2 个结构作为后续 WCNet 水体色度分类模型的基本结构单元。

4 水体色度分类方案构建

本章将设计一种水体色度分类方案,用于特定的水体色度的分类。方案由 2 部分内容组成,第 1 部分内容为水体图像预处理,第 2 部分内容为水体色度分类模型(WCNet 模型)的搭建,且模型以残差结构与 Inception-V3 结构为基本结构单元进行搭建。该水体色度分类方案可减少色度分类的运行时间,降低色度分类的迭代次数,提高色度分类的准确率。

4.1 图像预处理

水体图像的预处理主要包括图像灰度化、图像归一化、多特征分步边缘检测、ROI 区域提取以及图像数据增强等,具体细节在第 3 节已经阐述,图 11 为水体图像数据集预处理的示意图。

4.2 WCNet 模型搭建

4.2.1 WCNet 模型基本连接方式

图 12 为 WCNet 网络模型的基本模块示意图,在这个基本连接方式当中, x 为上一层的输出变量,WC_Inception(x)为改进优化的 Inception 模块输出

值, $Out(x)$ 为经过残差运算后的总的输出值,再使用 Relu 激活函数对总的输出值进行计算,把计算后的值当做总的输出,最后再输入下一层结构。

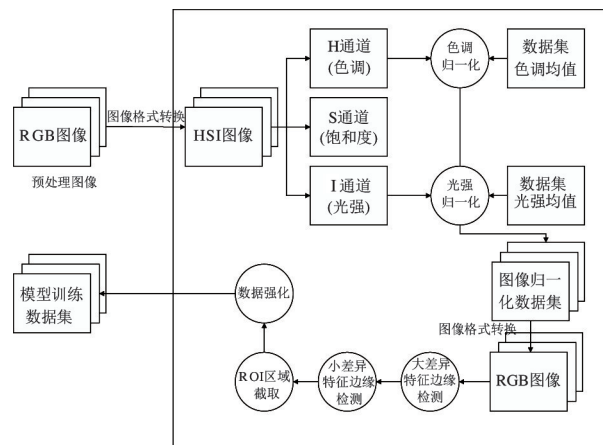


图 11 水体图像数据集预处理示意图

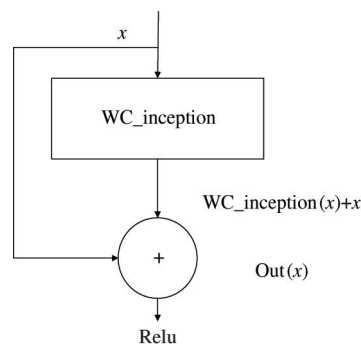


图 12 基本模块示意图

4.2.2 WCNet 15 和 WCNet 21 模型构建

由于实验采集到的水体图像都比较单一,且具备的特征点小,若在模型搭建过程中采用过多的层数,则极有可能导致过拟合现象以及模型退化等问题,所以需要设计一个浅层的 WCNet 模型。又因为不知道不同层数的 WCNet 网络对于水体色度识别效果有何影响,所以可再设计一个深层的 WCNet 模型。综上,在 WCNet 模型的设计方面,共设计了 2 种 WCNet 模型:15 层的浅层 WCNet 模型(WCNet 15),具体的神经网络示意图见图 13,以及 21 层的深层 WCNet 模型(WCNet 21),具体的神经网络示意图见图 14。

对于 WCNet 15 模型而言,其主要由图像输入端、Stem 结构、WC_Inception 结构、Dropout 结构、全连接层与分类层(softmax)结构 5 部分组成,图 13 为

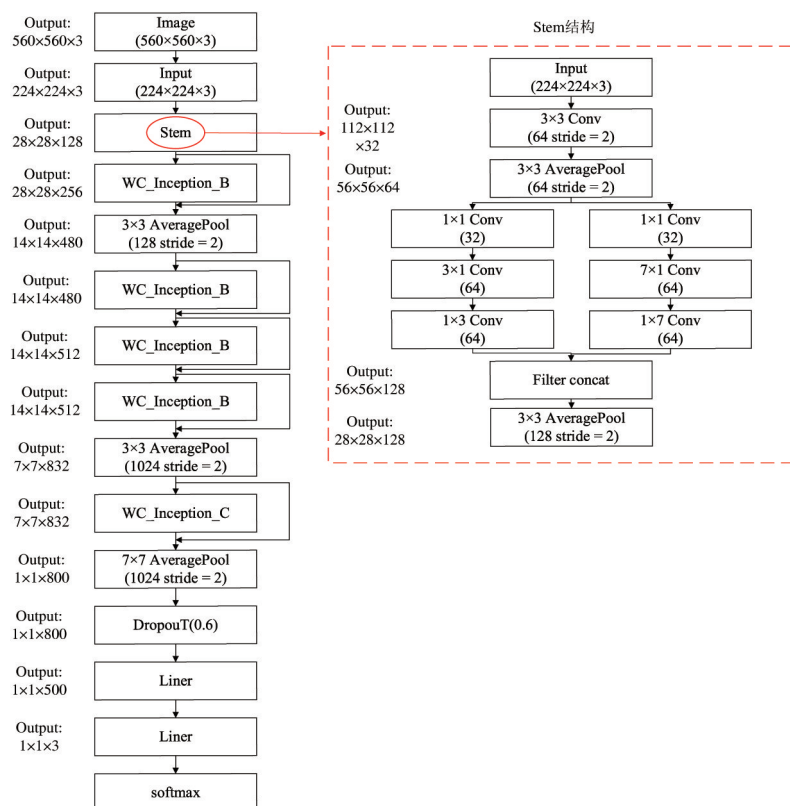


图 13 浅层 WCNet 15 神经网络示意图

具体结构示意图。

WCNet 15 模型各结构介绍如下。

(1) 图像输入端

对 560×560 像素的水样图像随机选择任意的 224×224 像素, 以此作为输入图像输入模型, 输入图像为 RGB 3 通道图像。

(2) Stem 结构

WCNet 15 中的 Stem 结构模块是图像输入端的直接连接模块, 模仿自 GoogleNet 模型, 其可对输入图像进行卷积操作。在 WCNet 15 的 Stem 结构中, 先采用 3×3 的卷积核图像进行卷积操作, 再采用均值池化层进行卷积层的均值池化操作。之后再进行一次并行的卷积操作, 并行线路 1 采用 1×1 、 3×1 和 1×3 的卷积核进行卷积操作, 并行线路 2 采用 1×1 、 7×1 和 1×7 的卷积核进行卷积操作。通过以上操作可以增加不同的卷积感受野以及使数据更加丰富, WCNet 15 中的 Stem 结构最终将输出 $28 \times 28 \times 128$ 的数据结构。

(3) WC_Inception 结构

WC_Inception 结构存在许多不同的模块结构,

而 WCNet 15 模型当中主要包含 2 种模块结构, 一种是 WC_Inception_B 结构, 另一种是 WC_Inception_C 结构, 这些结构将在 4.2.3 小节具体介绍。在 WCNet 15 模型当中, 共使用 4 次 WC_Inception_B 和 1 次 WC_Inception_C 操作, 模块结构之间采用残差结构进行连接, 目的是防止网络出现退化。其间还连接着 3×3 卷积核与 7×7 卷积核的均值池化层, 用于数据的下采样操作以及数据结构的转变工作, 具体连接见图 13。

(4) Dropout 结构

WCNet 15 模型当中设有 Dropout 结构, 且 Dropout 的初始参数设为 0.5, 目的是为了防止过拟合。

(5) 全连接层与 softmax 结构

在 Dropout 结构后面设置 2 个全连接层, 用于进一步连接卷积层的输出数据。2 个全连接层的设置使得数据可以平稳过度到最终的 3 个输出层。最终输出的数据为 $1 \times 1 \times 3$, 再使用 softmax 分类器进行水体色度分类。

以上便是 WCNet 15 模型的结构设计, 后为了探究不同网络深度以及不同感受野下的模型在水体

色度分类上的差异,再设计一个拥有 21 层的深层 WCNet 模型,即 WCNet 21 模型。

对于 WCNet 21 模型而言,自身层数的增加使得输入水体图像的尺寸也需要增大,输入水体图像的像素由 224×224 增大到 299×299 。此外为了增加 WCNet 21 模型的感受野,在模型中的 Stem 结构后加入 2 个 WC_Inception_A 模块结构,使得 WCNet 21 模型能获得更多的不同感受野下的水体图像特征。

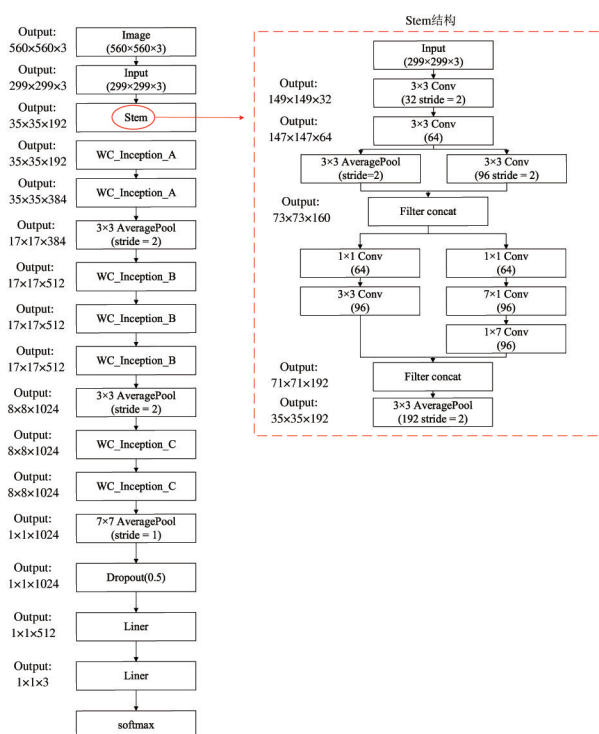


图 14 深层 WCNet 21 神经网络示意图

WCNet 21 模型各结构介绍如下。

(1) 图像输入端

相比于 WCNet 15, WCNet 21 的网络深度更深,所以输入图像应该更大,以保证后续的卷积操作可以顺利进行。所以对 560×560 像素的水样图像随机选择任意的 299×299 像素,以此作为输入图像输入模型,输入图像为 RGB 3 通道图像。

(2) Stem 结构

WCNet 21 模型中的 Stem 结构模块也是图像输入端的直接连接模块,其也可对输入图像进行卷积操作。在 WCNet 21 的 Stem 结构中,先采用 3×3 的卷积核对图像进行 2 次卷积操作,再采用均值池化层进行卷积层的均值池化操作。之后再进行一次并

行的卷积操作,并行线路 1 采用均值池化层进行池化操作,并行线路 2 采用 3×3 的卷积核进行卷积操作,后合并 2 条卷积线路。合并之后再进行一次并行的卷积操作,并行线路 1 采用 1×1 与 3×3 的卷积核进行卷积操作,并行线路 2 采用 1×1 、 7×1 和 7×1 的卷积核进行卷积操作。通过以上操作可以增加不同的卷积感受野以及使数据更加丰富,WCNet 21 中的 Stem 结构最终将输出 $35 \times 35 \times 192$ 的数据结构。

(3) WC_Inception 结构

对于 WCNet 21 模型,其主要包含 3 种模块结构,分别为 WC_Inception_A、WC_Inception_B 以及 WC_Inception_C,这些结构将在 4.2.3 小节具体介绍。WCNet 21 模型之所以使用多种 WC_Inception 模块结构,原因在于使用多种 WC_Inception 结构可以对不同尺寸的水体输入图像进行专门的卷积操作,防止不合适的卷积核进行卷积操作的情况发生,进而使得最终结果更准确。在 WCNet 21 模型当中共使用 2 次 WC_Inception_A、3 次 WC_Inception_B 和 2 次 WC_Inception_C 操作,其间使用均值池化层与残差结构加以连接,具体连接见图 14。

(4) Dropout 结构

WCNet 21 模型当中设有 Dropout 结构,且 Dropout 的初始参数设为 0.5,目的是为了防止过拟合。

(5) 全连接层与 softmax 结构

在 Dropout 结构后面设置 2 个全连接层,用于进一步连接卷积层的输出数据。2 个全连接层的设置使得数据可以平稳过渡到最终的 3 个输出层。最终输出的数据为 $1 \times 1 \times 3$,再使用 softmax 分类器进行水体色度分类。

4.2.3 WC_Inception 模块设计

WC_Inception 模块结构源于 GoogleNet 模型当中的 Inception 模块结构,由 Inception 模块结构调整所得。WC_Inception 模块结构有以下优点:(1) 模块步长均为 1,操作时不会使得图像数据的尺寸发生变化,避免图像数据缩小过快现象;(2) 由于模块自身的串联结构,使得搭建模型更加简便;(3) WC_Inception 模块种类多,不同的 WC_Inception

模块结构可以对不同大小的输入图像以及不同阶段的输入图像进行操作。常见的 WC_Inception 模块结构有 WC_Inception_A 模块、WC_Inception_B 模块以及 WC_Inception_C 模块, 以下为各模块结构的相关介绍。

WC_Inception_A 模块结构图如图 15 所示, 该模块有 4 个分支。最左边的分支为 1×1 卷积核的卷积操作, 其作用是为了使参数量降低, 同时使非线性关系得以增加。最右边的分支为 1×1 、 1×7 与 7×1 卷积核的串联卷积运算, 其作用是为了使卷积运算视野增大, 使大面积特征的提取工作得以增加, 避免水体图像中小面积的干扰因素带来影响。4 个分支进行并联的卷积运算, 使得结构模块的鲁棒性以及非线性得以增强, 使得水体图像中的更多特征数据能被加以提取。

WC_Inception_A 模块只用于 WCNet 21 模型, 且只用于该模型的前部分。这是因为 WCNet 21 模型前期的数据量较大, 为 $35 \times 35 \times 3$ 。由于数据量大, 所以大型卷积核需要被使用, 大型卷积核的使用可以更好地提取大面积上的数据, 同时使局部数据的权重得以降低, 这样设计与操作使得 WCNet 21 模型初期的鲁棒性得到了很好的提高。

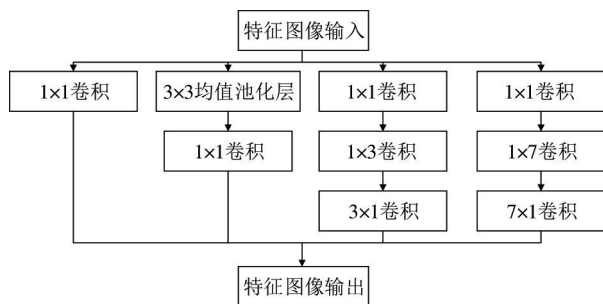


图 15 WC_Inception_A 结构图

WC_Inception_B 模块结构图如图 16 所示, 该模块有 4 个分支, 最右边的分支为 5 个卷积核的串联卷积操作、中间分支为池化以及卷积操作等。

WC_Inception_B 模块主要用于 WCNet 15 以及 WCNet 21 模型的前中部分。这是因为 WCNet 15 以及 WCNet 21 模型前中期的数据量较小, 分别为 $14 \times 14 \times 3$ 与 $17 \times 17 \times 3$ 。由于数据量小, 所以小型卷积核需要被使用, 小型卷积核的使用使得当数据

量较小时, 模型依旧可以进行多次卷积操作, 进而使得水体图像中的更多特征数据能被加以提取。这样设计与操作使得 WCNet 15 与 WCNet 21 模型的前中期特征提取效果得以加强, 同时也丰富了水体数据特征。

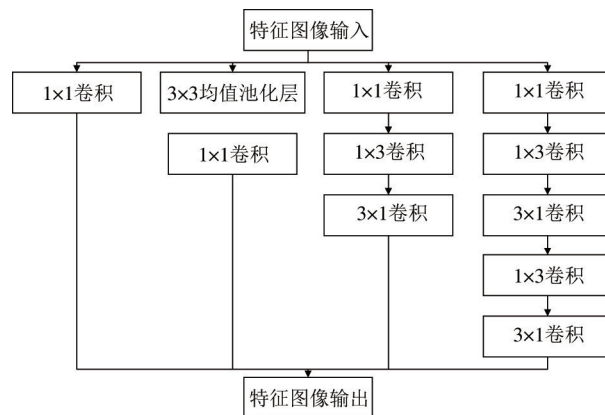


图 16 WC_Inception_B 结构图

WC_Inception_C 模块结构图如图 17 所示, 该模块有 3 个分支, 最右边的分支为 3 个卷积核的串联卷积操作, 中间分支进行卷积以及池化等操作。

WC_Inception_C 模块主要用于 WCNet 15 以及 WCNet 21 模型的后部分。这是因为 WCNet 15 以及 WCNet 21 模型后期的数据量较小, 分别为 $7 \times 7 \times 3$ 与 $8 \times 8 \times 3$ 。且 WC_Inception_C 模块能处理好较小的水体图像数据, 对其能进行进一步的有效提取, 同时增强 WCNet 15 以及 WCNet 21 模型的非线性。

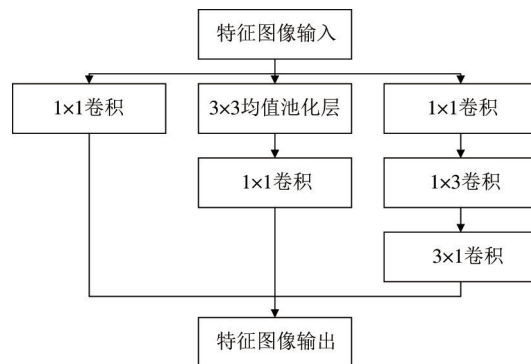


图 17 WC_Inception_C 结构图

4.2.4 其他设计

针对 WCNet 15 与 WCNet 21 模型, 还需进行池化层、激活函数、损失函数等方面的设置。

池化层:使用均值池化层作为 2 种 WCNet 模型的池化层,使用均值池化层的原因在于获得的水体图像是存在渐变的,其间会有一些干扰像素干扰水体色度的分类。而使用均值池化层可使得水体图像的信息被更好地了解,进而避免上述问题的发生。

激活函数:使用 Relu 函数作为 2 种 WCNet 模型的激活函数。使用 Relu 函数的原因在于其分类能力较强,可防止过拟合。

正则化:当数据经过 Relu 激活函数处理后,再进行 Batch Normalization 正则化操作。

损失函数:使用交叉熵损失函数作为 2 种 WCNet 模型的损失函数,但该损失函数可能会因为标签的分布不均而产生过拟合。为了防止过拟合现象的发生,将标签平滑正则化(label smoothing regularization, LSR)方法应用到损失函数当中去。

各神经网络的参数数量如表 4 所示,且参数数量占比以 VGG 16 为标准。就网络层数而言,表中模型都比较接近,但 WCNet 15 以及 WCNet 21 模型的参数数量却十分低,占比只有 1.5% 与 3.3%。由

此证明了在不同分类阶段采用不同的卷积策略可以减少模型自身的参数,而参数数量的减少可使得各个模型训练速度加快,减少训练时间,当运用到实际检测中时,可做到实时检测。

表 4 各神经网络参数表

神经网络名称	参数数量/ 10^3	占比
VGG 16	138 357	100%
GoogleNet _ V1	6804	4.9%
GoogleNet _ V3	24 734	17.9%
ResNet 18	33 161	24.0%
WCNet 15	2110	1.5%
WCNet 21	4607	3.3%

5 实验与结果分析

5.1 实验环节

表 5 为 2 个 WCNet 网络模型训练参数的相关设置,设置以后进行模型的训练工作。

表 5 模型参数设置

模型参数	备注	网络名称	
		WCNet 15	WCNet 21
num _ classes	分类数	3	3
input _ size	输入尺寸	224×224	299×299
data _ type	数据类型	ImageFolder	ImageFolder
Dropout		0.5	0.5
lr _ ori	初始学习率	0.01	0.01
lr _ config	学习率衰减	100epoch, lr _ ori $\times$ 0.5	100epoch, lr _ ori $\times$ 0.5
optimizer	优化器	Adam($\beta_1 = 0.9$, $\beta_2 = 0.99$)	Adam($\beta_1 = 0.9$, $\beta_2 = 0.99$)
batch _ size	批量数量	8	8
epoch	迭代次数	5000	5000
shuffle	随机打乱	TRUE	TRUE
loss _ f	损失函数	LSR	LSR

模型训练主要包含以下 6 个步骤。

- (1) 对模型权重进行随机初始化操作;
- (2) 对模型参数进行前向传播;
- (3) 对模型损失函数进行计算;
- (4) 对模型参数进行反向传播;
- (5) 使用梯度优化器,优化与更新所有模型的参数;

(6) 将最优的模型参数进行保存。

WCNet 15 模型的准确率变化如图 18 所示。由变化图易知,WCNet 15 模型在整个迭代过程中收敛比较迅速,且准确率在约 2000 次迭代之后趋于稳定。模型的最优准确率为 96.3% 左右。

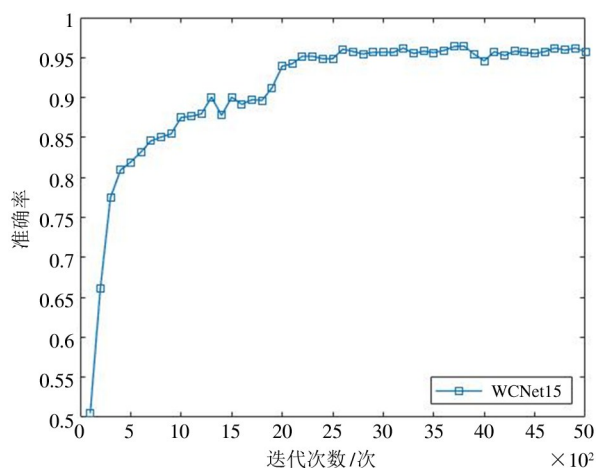


图 18 WCNet 15 网络准确率变化图

WCNet 21 模型的准确率变化如图 19 所示。由变化图易知,该模型收敛也比较迅速,准确率平稳于 2000 次迭代之后。模型的最优准确率为 97.8% 左右。

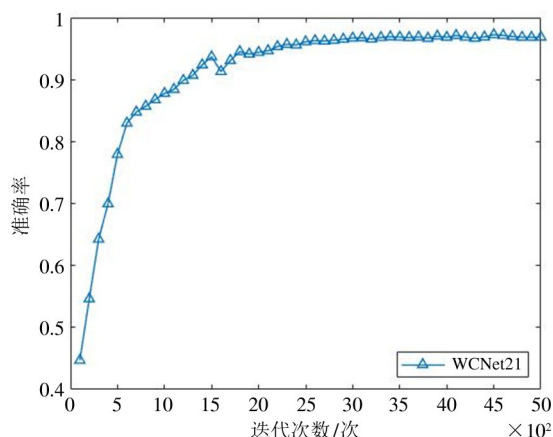


图 19 WCNet 21 网络准确率变化图

由图 18 以及图 19 可知,WCNet 15 与 WCNet 21 模型具有较高的水体色度分类准确率,能进行相关的水体色度识别工作。

5.2 结果分析与模型比较

为了更清晰地进行 GoogleNet-V3、ResNet 18、WCNet 15 以及 WCNet 21 4 种模型的水体色度分类准确率对比,绘制各模型准确率变化趋势图,如图 20 所示。

表 6 为 GoogleNet-V3、ResNet 18、WCNet 15 以及 WCNet 21 4 种模型水体色度分类的最优准确率。

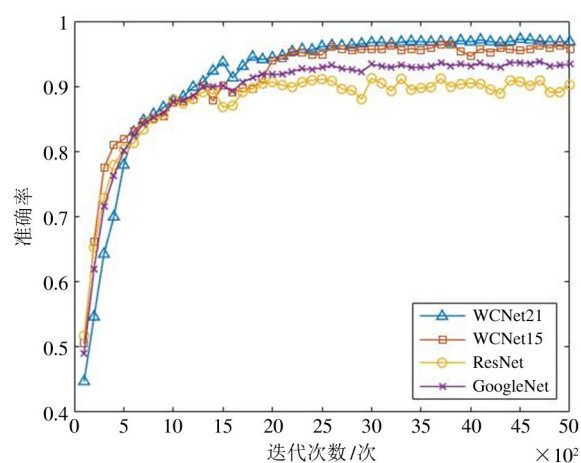


图 20 各模型准确率变化趋势图

表 6 模型最优准确率表

模型类型	Google Net-V3	ResNet 18	WCNet 21	WCNet 15
最优准确率	93.8%	90.6%	97.8%	96.3%

4 个模型的实验条件与装置均相同,由图 20 与表 6 的数据分析可知,WCNet 21 模型分类准确率最高,表现最好,可作为最终的水体色度分类模型进行使用。

6 结 论

本研究建立了基于深度学习下的水体色度分类方案,包括无人机拍摄获取输入数据、图像预处理以及 WCNet 神经网络的搭建等,从最终的训练结果得出以下结论。

(1)模型分类的准确率大小与神经网络深度有着正相关的关系,采用了更深层网络的 WCNet 21 模型分类准确率要高于浅层的 WCNet 15 模型。

(2)Inception 结构能提升水体色度分类的准确率,使用了 Inception 结构的 GoogleNet-V3、WCNet 15 以及 WCNet 21 3 个模型分类准确率均高于未使用 Inception 结构的 ResNet 18 模型。且 WC_Inception 结构对分类准确率的提高更为明显,其是以 Inception 结构为基础,后按实际需求进行结构调整与优化的新模块结构。

(3)残差结构能提升水体色度分类的准确率,使用了残差结构的 WCNet 15 与 WCNet 21 模型的

分类准确率要高于未使用残差结构的 GoogleNet-V3 模型。

(4) WCNet 模型的分类准确率要高于 GoogleNet-V3 与 ResNet 18 模型,且平稳后的波动更小,稳定性更佳。而在 WCNet 模型当中,WCNet 21 模型要优于 WCNet 15 模型。

虽然本文设计的水体色度分类模型有一定的优势,但是该模型也存在一定的局限性。首先,由于水体图像样本数量上的局限,使得只能在某一色度范围内进行有效的检测,不能进行更加细化的分组。其次,虽然设置了多特征分步边缘检测方法用于水体图像的边缘检测,但仍需人为地进行参数调整。最后,对于在较差环境条件下采集到的水体图像,很难用设计的图像预处理进行数据提取,进而使得整个水体分类工作无法进行。

下一阶段研究将从以下 2 个方面入手,一是不断完善水体数据库,二是加大探索图像预处理的手段。进一步提升模型的色度检测精度与检测范围,降低水体色度分类模型对环境条件的要求。

参考文献

- [1] 赵晓伟,沈为民,黄杰,等. 基于三波长透射率的水体色度检测标准[J]. 环境工程学报, 2013, 7(12): 4766-4772.
- [2] KAO C M, CHOU M S, FANG W L, et al. Regulating colored textile wastewater by 3/31 wavelength ADMI methods in Taiwan[J]. Chemosphere, 2001, 44(5): 1055-1063.
- [3] 王安,谢宇,梁柱. 水中色度测定的研究[J]. 中国环境监测, 2000, 16(2): 37-40.
- [4] 刘春,郭春娟. 用分光光度计测定地面水色度[J]. 环境科技(辽宁), 1994, 1: 75-76.
- [5] GRABAS M. Organic matter removal from meat processing wastewater using moving bed biofilm reactors[J]. Environment Protection Engineering, 2000, 26(1): 55-62.
- [6] YU R F, CHEN H W, CHENG W P, et al. Measurements of wastewater true color by 4/6 wavelength methods and artificial neural network[J]. Environmental Monitoring and Assessment, 2006, 118(1-3): 195-209.
- [7] 曾凡亮,罗先桃. 分光光度法测定水样的色度[J]. 工业水处理, 2006, 9: 69-72.
- [8] 孙自瑾,张因,高春光,等. 分光光度法测定聚四氢呋喃的色度[J]. 山西大学学报(自然科学版), 2009, 32(3): 447-450.
- [9] 朱建华. 分光光度计测量 2 类水体光谱吸收系数的关键技术[J]. 海洋技术, 2003, 1: 34-39.
- [10] 巫涛江. 基于光谱法的水质色度在线检测方法研究[D]. 重庆: 重庆理工大学电气与电子工程学院, 2020: 2-3.
- [11] MEMON S F, ALI M M, PEMBROKE J T, et al. Measurement of ultralow level bioethanol concentration for production using evanescent wave based optical fiber sensor[J]. IEEE Transactions on Instrumentation and Measurement, 2018, 67(4): 780-788.
- [12] 黄杰,沈为民,楼俊,等. 基于光谱及色差计算测定水质的色度[J]. 光谱实验室, 2012, 29(6): 3399-3403.
- [13] HU X H, MEGRET P, CAUCHETEUR C. Surface plasmon excitation at near-infrared wavelengths in polymer optical fibers[J]. Optics Letters, 2015, 40(17): 3998-4001.
- [14] 李文,杨守波,罗学科. 基于远程图像色度的点源水质监测方法[J]. 环境工程学报, 2015, 9(9): 4603-4608.
- [15] 何晓群. 关于 6 Sigma 与 3 Sigma 的比较[J]. 数理统计与管理, 2006, 2: 175-177.
- [16] 李庆忠,刘洋. 基于改进 Canny 算子的图像弱边缘检测算法[J]. 计算机应用研究, 2020, 37(S1): 361-363.
- [17] 田垅,刘宗田. 最小二乘法分段直线拟合[J]. 计算机科学, 2012, 39(S1): 482-484.
- [18] 马立新,豆晨飞,宋晨灿,等. 基于 VGGNet 算法的绝缘子无损检测[J]. 光电工程, 2021, 48(1): 38-45.
- [19] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolution[C] // Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 1-9.
- [20] 余琼芳,黄高路,杨艺,等. 基于 AlexNet 深度学习网络的串联故障电弧检测方法[J]. 电子测量与仪器学报, 2019, 33(3): 145-152.
- [21] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C] // Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.

Research on water color classification algorithm based on deep learning

CHEN Yong, CHEN Jian, CHEN Yi, PEI Zhi, YI Wenchao, WANG Cheng, ZHANG Wenzhu

(College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023)

Abstract

Deep learning technology is used for non-contact, fast water color detection and classification. To establish the data set, the unmanned aerial vehicle is applied to collect the water images and the colorimeter is used to classify the calibrated images. Then images are normalized to reduce the impact of environmental factors on the classification results. A multi-feature step-by-step edge detection algorithm is designed to detect the edge of the water image and eliminate irrelevant pixels. The proposed algorithm applies VGG 16, Googlenet-V3 and Resnet 18 convolutional neural networks to construct and train the water color classification model, and selects the concept structure and residual structure as the basic building units of WCNet 15 and WCNet 21 neural network models specifically for water color classification. The model parameters are trained on the training set, and the accuracy of the two models is compared by using the validation set. WCNet 21 is more suitable for water color classification as it has higher accuracy. The optimal accuracy of model WCNet 21 can reach 97.8%, which meets the standards of water color classification and can be applied to real-life water color classification.

Key words: water quality detection, chromaticity classification, deep learning, edge detection, neural network