

DOI:10.19651/j.cnki.emt.2208869

## 基于深度神经网络的液体视觉识别研究\*

钟扬<sup>1</sup> 吴黎明<sup>1</sup> 温腾腾<sup>1</sup> 伍冠楚<sup>1,2</sup> 王桂棠<sup>1,2</sup>

(1. 广东工业大学机电工程学院 广州 510006; 2. 佛山沧科智能科技有限公司 佛山 528225)

**摘要:** 针对液体表面特征少,区分度低,机器视觉难以有效识别检测的问题,通过使用两束不同波长的激光光源同时照射液体来提高不同液体之间的区分度,设计了数据集自动采集装置为模型训练提供了大量有效的样本,并构建了基于 EfficientNetV2 深度神经网络的视觉识别模型,模型引入 cosine 学习率衰减,调节获得最佳超参数后,形成最优方式实现高效训练,进一步提升了预测精度,结果表明视觉检测系统能够获得 100% 的测试准确率,成功解决了液体视觉检测中特征少的难题。

**关键词:** 机器视觉;激光;深度神经网络;液体视觉检测

**中图分类号:** TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.2040

## Research on liquid vision recognition based on deep neural network

Zhong Yang<sup>1</sup> Wu Liming<sup>1</sup> Wen Tengting<sup>1</sup> Wu Guanchu<sup>1,2</sup> Wang Guitang<sup>1,2</sup>

(1. School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China;

2. Foshan Cangke Intelligent Technology Co., Ltd., Foshan 528225, China)

**Abstract:** Aiming at the problem that liquid surface features are few and discrimination is low, which is difficult to be recognized and detected effectively by machine vision, two laser light sources with different wavelengths are used to irradiate liquid at the same time to improve the discrimination between different liquids. An automatic collection device of data set is designed to provide a large number of effective samples for model training and then a visual recognition model based on EfficientNetV2 deep neural network was constructed. After introducing cosine learning rate decay into the model and regulating super-parameters to the best, the optimal method was formed to realize efficient training, and the prediction accuracy was further improved. The results showed that the visual detection system could obtain 100% prediction accuracy, and successfully solved the problem of few features in liquid visual detection.

**Keywords:** machine vision perception; laser; deep neural network; liquid visual detection

## 0 引言

随着人工智能技术的广泛应用,需要对事物进行智能感知,液体识别方法是一个值得研究的方向。文献[1]提出利用光线传感器检测液滴的光纤中光信号变化规律来检测液体,方法能够准确区分相同颜色液体但是检测时间长。文献[2-3]提出使用 X 射线透射生成图像,然后再对图像进行处理的方法对液体进行识别,这种方法能有效区分液体但生产成本低。文献[4]提出使用电子舌对有味道液体信息进行采集,然后使用神经网络进行识别,这种方法准确率高但局限于有味道液体。文献[5]提出使用摄像头对液体散斑图信息进行采集,然后使用卷积神经网络对这些信息进行识别;文献[6]提出使用摄像头对液体色度进行测量的方法检测水质;文献[7]提出使用对激光散射的图像进行处

理来检测城市颗粒物;这 3 种方法能有效区分颜色相近对象,但结构非常复杂实现难度大。文献[8]提出使用由光源、相机以及软件组成的机器视觉系统来对滴定后液体颜色进行识别,这种方法识别准确率高但要求液体必须有颜色区分。

文献[9]提出机器视觉能够应用于液体检测中,并证明其具有快速的优势。然而,颜色相近液体的视觉特征区分度低,例如,对于牛奶、豆奶、酸奶等白色液体以及可乐、生抽、老抽等黑色液体,使用光照自动增益相机拍摄的图像特征难以区分,而取消自动增益设置后,识别系统会因为受到环境光的影响加大而导致鲁棒性差。为减少环境光的影响而使用外部可控光源,则会进一步产生反光、光斑等问题,严重影响特征区分度。对于人的肉眼难以区分的透明液体

收稿日期:2022-01-17

\* 基金项目:国家自然科学基金(61705045)、佛山广工大研究院创新创业人才团队计划(20191108)项目资助

如纯净水、糖水、盐水、白醋等,即便调整光照、设置相机参数或采用暗室拍照等操作,使用机器视觉方法识别效果不佳。

不同液体介质的吸收率、折射率、反射率、色散等性质不同,通过激光照射会形成不同的视觉特征。

本文根据液体介质的性质,创新性地提出了一种基于激光照射实现机器视觉液体识别检测的方法,并通过搭建端到端的基于深度神经网络的识别模型,对不同学习率策略进行调优,相对于其他模型提高了识别准确率并降低了交叉熵损失,为难区分液体的视觉识别检测提供了一条新路径。

## 1 基于机器视觉的液体识别模型

EfficientNetV2 网络模型是在 2021 年由 Tan 等<sup>[10]</sup>发布出来的,它在 CIFAR、Cars、Flowers 等公共数据集上显著优于其他网络模型,同时在 ImageNet 上达到 87.3% 的最高精度,以 2.0% 精度优于 ViT,且训练速度快 5~10 倍,所以本文选择 EfficientNetV2 作为基于机器视觉的液体识别模型<sup>[10-16]</sup>。EfficientNetV2 网络中有 4 个重要部分,分别是 SE 网络(squeeze-and-excitation networks, SEN)、随机深度网络(stochastic depth network, SDN)、MBConv 网络与 Fused-MBConv 网络。

### 1) SE 网络

如图 1 所示,SE 网络分为两个部分,分别是压缩操作以及激励操作。压缩操作实现上使用的是一个全局平均池化(global average pooling, GAP),将形状(shape)为  $W \times H \times C$  的输入特征图压缩成为一个  $1 \times 1 \times C$  的特征图。激励操作则将压缩输出的  $1 \times 1 \times C$  特征图经过两个全连接层(fully-connected layer, FC),第 1 个全连接层使用 SiLU 激活函数,第 2 个全连接层使用 sigmoid 激活函数。最后,将激励操作输出的特征图与原有特征图在通道方向上进行元素级乘积计算。SE 网络学习了各个通道之间的相关性,虽然增加了一定的计算量,但筛选得到通道注意力,提升了检测精度。

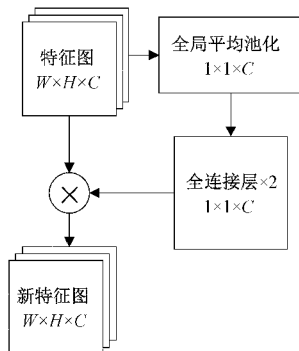


图 1 SE 网络结构图

### 2) 随机深度网络

随机深度网络相对于普通的 ResNet 网络来说,区别是

它在训练过程中随机地丢弃一些网络层,能够一定程度上减小在训练时的网络深度,加快训练速度,同时在预测的时候保持原有的网络深度以保证网络的表达能力。在 EfficientNetV2 引入的 Dropout 机制是基于随机深度网络的。

如图 2 所示,随机深度网络实现方法是在训练过程中,先对上一层的输入  $H_{i-1}$  进行  $f_i$  映射(映射即为残差块的两层卷积),然后用这个映射结果  $f_i(H_{i-1})$  乘以一个满足伯努利分布的变量  $b_i$ ,然后再与上一层网络的输出  $id(H_{i-1})$  进行元素相加,最后使用 ReLU 激活函数进行激活。当  $b_i = 1$  时,整个网络结构相当于一个残差块,而当  $b_i = 0$  时,残差块没有加入到网络中,整个网络结构退化成一个恒等函数加上 ReLU 激活函数的映射  $ReLU(id(H_{i-1}))$ 。

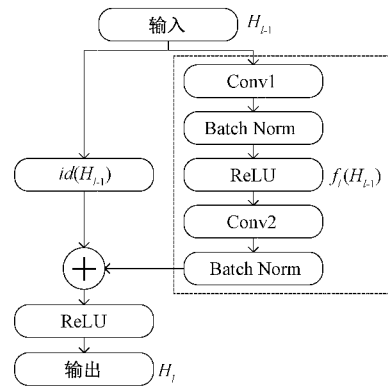


图 2 随机深度网络结构

整个过程可以表示为:

$$H_i = \text{ReLU}(b_i f_i(H_{i-1}) + id(H_{i-1})) \quad (1)$$

### 3) MBConv 和 Fused-MBConv 网络

如图 3 左半部分所示,MBConv 网络模块源于 MobileNetV3 网络中的逆残差块(inverted residual block, IRB),而 EfficientNet 将 MBConv 激活函数替换成 SiLU 激活函数,并且每个 MBConv 模块均加入了 SE 网络模块。

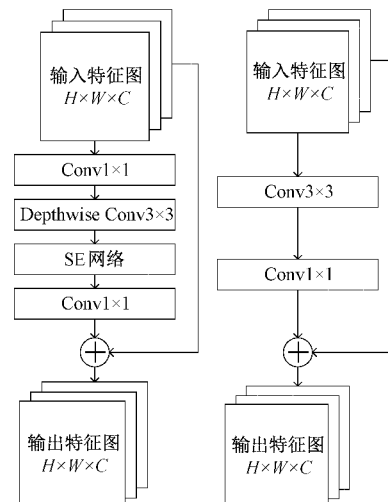


图 3 MBConv 与 Fused-MBConv 结构

如图 3 右半部分所示, EfficientNetV2 通过在网络的浅层(1~3 层)使用 Fused-MBConv 来代替 MBConv 以解决 depthwise 卷积无法充分利用网络导致的训练速度较慢的问题, 并删除了 SE 网络模块来提升模型效果。

## 2 实验环境搭建及数据集采集

### 2.1 视觉识别平台搭建

视觉识别平台用于获取液体的图像信息, 如图 4 所示, 识别平台包括铝制龙门架、光电开关控制的激光模块、工业相机、暗室箱、数据集采集硬件平台。其中, 龙门架用于装配工业相机; 激光模块包含两种不同波长的激光, 可以实现两束激光同时照射到液体中, 形成不同波长下的视觉特征; 暗室箱保证了识别环境不受外界光照干扰, 提高图像成像效果, 同时提高鲁棒性。系统中入射激光的波长分别为 405 和 532 nm。

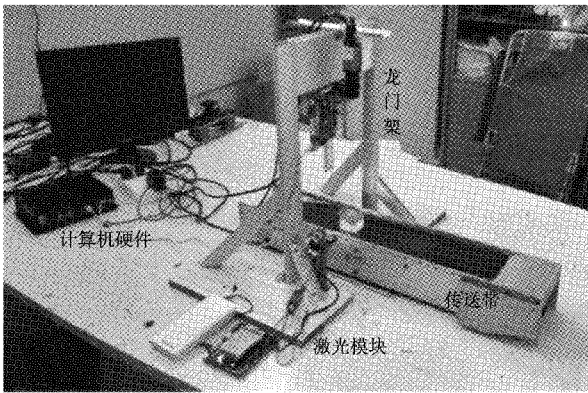


图 4 视觉识别平台

### 2.2 数据集采集硬件平台

在机器视觉识别系统中, 建立分类模型所需的数据集是十分重要的一个环节, 但是人工采集有着过程繁琐、用时长、采集工作重复度高等问题。为了提升采集图像数据的效率, 设计了自动化移动旋转采集装置, 解决了人工采集费时费力的问题。

如图 5 所示, 装置总体由滑托和滑槽构成, 滑托由两个轮组控制的传送带带动, 控制着滑托直线运动, 两个轮组均由舵机带动, 滑托上面的平台部分也由舵机控制, 控制着平台顺时针或逆时针旋转。3 个舵机(两个控制滑托移动, 一个控制平台旋转)均由 STM32F103RCT6 最小系统板控制, 舵机的信号线接 STM32F103RCT6 产生的 PWM 脉冲, 而另外两根线分别接 5 V 和 GND, 基于 PWM 脉宽调制控制舵机旋转的角度。通过纵向、横向以及旋转方向改变样品的位置, 可以获得具有不同位置偏移的图像, 为训练深度学习模型提供了大量有效的样本。

### 2.3 实验对象分类

为验证模型的有效性, 实验选用了 20 种液体, 如表 1 所示。

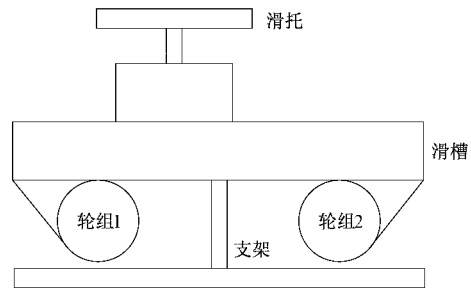


图 5 数据集采集硬件平台

表 1 实验对象表

| 编号 | 中文名                  | English code | 液体颜色  |
|----|----------------------|--------------|-------|
| 0  | 盐水<br>(0.9% NaCl 溶液) | brine        |       |
| 1  | 雪碧                   | sprite       | 透明    |
| 2  | 糖水<br>(0.3% 白砂糖溶液)   | sugar water  |       |
| 3  | 白醋                   | vinegar      |       |
| 4  | 纯净水                  | water        |       |
| 5  | 牛奶                   | milk         |       |
| 6  | 豆奶                   | soy milk     | 白色    |
| 7  | 酸奶                   | yogurt       |       |
| 8  | 可乐                   | cola         |       |
| 9  | 老抽                   | laochou      | 黑色    |
| 10 | 生抽                   | shengchou    |       |
| 11 | 苹果汁                  | apple juice  |       |
| 12 | 啤酒                   | beer         | 淡黄色   |
| 13 | 花生油                  | peanut oil   |       |
| 14 | 白葡萄酒                 | white wine   |       |
| 15 | 鸡汤(鸡精溶液)             | chicken soup | 浑浊状黄色 |
| 16 | 香油                   | sesame oil   | 棕色    |
| 17 | 乌龙茶                  | tea          |       |
| 18 | 咖啡                   | coffee       | 黑色    |
| 19 | 橙汁                   | orange juice | 深黄色   |

### 2.4 数据集采集

设计并采用数据集采集硬件平台, 获取不同角度、位移偏差的样本图像, 提高数据样本的类型广度, 有助于提高识别模型的泛化能力。

#### 1) 训练集与验证集采集

由于实验中涉及颜色相近液体(包括完全透明液体), 所以进行数据集采集以及最终测试的时候, 需要使得不同类别液体间的图片特征信息要有明显不同, 同时相同类别液体间的图片特征信息一致性足够大, 才能保证最终分类的准确性以及鲁棒性。为确保模型能够充分的学习到图像中的关键特征信息而非液体所在位置信息, 实验过程中分别采用了移动 y 轴方向、移动 x 轴方向、绕 z 轴顺时针及

逆时针旋转的多方向数据集采集装置移动方法来获取多角度液体信息。其中  $y$  轴方向为传送带的运动方向,  $x$  轴方向为与传送带同平面且垂直于  $y$  轴的方向, 而  $z$  轴方向为垂直于  $x$ 、 $y$  轴相交所成平面的方向。通过在  $x$  轴方向移动液体 20 次,  $y$  轴方向移动液体 3 次并在绕  $z$  轴旋转方向旋转 3 液体次操作, 最终每一类获取了 180 张图片, 共 3 600 张图片。所采集的数据按 9 : 1 分别划分为训练集和验证集, 采样均为随机采样。

上述步骤划分出来的训练集用于模型训练, 而验证集则用于模型训练过程中的超参数调优以及观测模型是否存在过拟合的现象。

## 2) 测试集采集

为了在最终阶段准确地测试模型并分析最终的实验结果, 在真实的实验环境下, 人工采集了与验证集不同的 360 张图片作为测试集。

## 2.5 原始图像分析

对实验环境下采集的图片以及对使用数据集采集装置采集的数据集进行分析, 根据目视观察可以将较难区分的液体分为 4 组: 第 1 组为盐水、雪碧、糖水、白醋和纯净水等透明液体; 第 2 组为牛奶、豆奶和酸奶等白色液体; 第 3 组为老抽、生抽和可乐等黑色液体; 第 4 组为苹果汁、啤酒、花生油和白葡萄酒等淡黄色液体, 图片中 532 nm 激光照射左半边, 405 nm 激光照射右半边。

不同波长的激光照射不同液体介质时, 因其吸收率、折射率、反射率、色散等性质不同, 会形成不同的视觉特征。因为使用 3 种及以上的波长对液体进行照射控制难度较大且实验时会产生更多的变量, 同时效果提升并不明显, 所以经测试选择使用 532 nm 和 405 nm 激光时是最适合的。

通常, 雪碧在取样时会产生丰富的气泡, 辨识度高, 而其他透明液体如纯净水、盐水、糖水、白醋等, 肉眼无法辨识。图 6 为暗室中通过双激光照射得到的图像, 如图 6(a) 所示, 在 0.9% 盐水中, 杯壁产生的光斑比纯净水大且能够观测到一条光束; 如图 6(b) 所示, 可看到激光在穿过产生气泡的雪碧所呈现图像为多处发光圆圈; 如图 6(c) 所示, 照射糖水时, 除了杯壁产生光斑外, 产生一条非常亮的光束; 如图 6(d) 所示, 照射白醋时则会产生几个光斑和传输光束, 在光束附近会有很多小点并且产生光斑在远离激光处杯壁会产生一条横线, 同时相较于盐水其右半边会产生一条更加明显的传输光束; 如图 6(e) 所示, 在纯净水中只在杯壁产生几个光点, 观察不到传输光束。

白色液体组分别是牛奶、豆奶、酸奶, 在人的肉眼下虽然能够区分但是在工业相机的拍摄下辨识度很低。图 7 为白色液体组各个液体在暗室通过波长 405 nm 和 532 nm 的激光照射产生的图像。激光照射时, 无法观察到波长为 405 nm 的激光器在白色液体组中的传输光束, 但是波长为 532 nm 的激光在穿过白色液体组各个液体时均产生了一个显著的大光斑, 如图 7(a) 所示, 牛奶的光斑与黑影具有

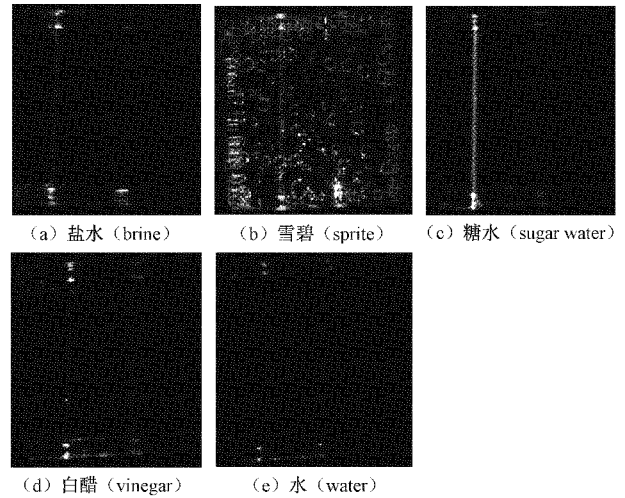


图 6 透明液体在暗室激光下图像

较为明显的轮廓; 如图 7(b) 所示, 豆奶的圆形光斑则较小, 且光斑与黑影有过渡区域, 仍然能够较为清晰地看到光斑的轮廓; 如图 7(c) 所示, 激光照射酸奶产生的光斑无明显轮廓。

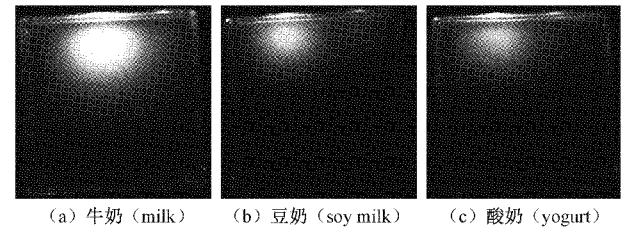


图 7 白色液体在暗室激光下图像

图 8 为黑色物体组的各个液体在暗室中通过双激光照射产生的图像。如图 8(a) 所示, 在可乐的图像上, 可以看到在左侧激光照射下产生了一条不清晰的光线段, 右侧激光照射下产生一个光点和光线; 如图 8(b) 所示, 老抽在左侧激光照射下产生了一个光点, 右侧激光照射下也产生一个光点; 如图 8(c) 所示, 在生抽图像中, 在 532 nm 激光的照射下产生了一个明显的光点和一束短促的光线段。

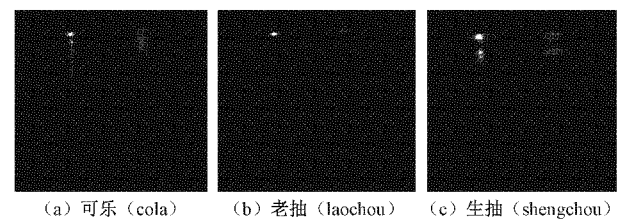


图 8 黑色液体在暗室激光下图像

图 9 为淡黄色液体组各个液体在暗室中通过双激光照射产生的图像。苹果汁在左侧激光照射下能够产生几个亮光点和一条细传输光线, 在右侧激光照射下产生光点以及一条粗传输光线; 啤酒在左侧激光照射下产生一条亮于苹果汁的传输光线, 右侧激光产生一条明显亮于苹果汁的一

条传输光线以及两个光点；花生油在左侧激光的照射下产生一条明显亮于苹果汁以及啤酒的传输光线以及几个光点并且图像底部产生一条水平的光线段，在右侧激光照射下能够产生明显亮于苹果汁以及啤酒的亮传输光线，并且光线从上到下慢慢消失；白葡萄酒能够在左侧激光照射下产生几个光点以及一条非常淡的传输光线，在右侧激光的照射下能够产生一条较暗的传输光线以及几个光点。

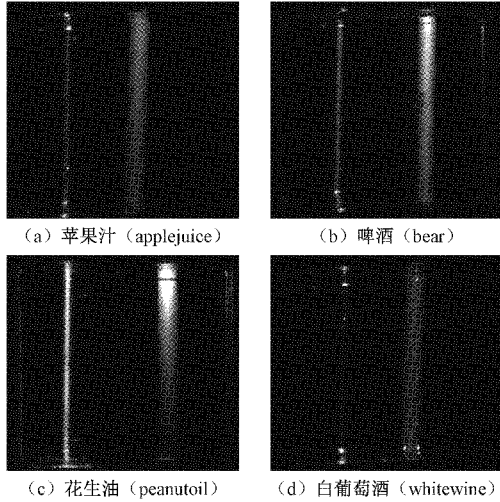


图 9 淡黄色液体在暗室激光下图像

### 3 视觉识别模型搭建及训练

#### 3.1 视觉识别模型搭建

研究采用了 EfficientNetV2-S 网络框架设计视觉识别模型。如表 2 所示, EfficientNetV2-S 中 SE 表示使用了 SE 模块, 而 0.25 表示 SE 模块的第 1 个全连接层神经元的个数是输入这个 MBCConv 模块的矩阵其通道数的 0.25 倍。layers 代表的 Operator 内的网络模块重复堆叠次数, 而 Stride 表示第 1 个堆叠中卷积核的步长。

表 2 EfficientNetV2-S 网络结构

| Stage  | Operator              | Stride | Channels | layers |
|--------|-----------------------|--------|----------|--------|
| Input  | —                     | —      | 3        | —      |
| 0      | Conv3x3               | 2      | 24       | 1      |
| 1      | Fused-MBConv1, k3x3   | 1      | 24       | 2      |
| 2      | Fused-MBConv4, k3x3   | 2      | 48       | 4      |
| 3      | Fused-MBConv4, k3x3   | 2      | 64       | 4      |
| 4      | MBConv4, k3x3, SE0.25 | 2      | 128      | 6      |
| 5      | MBConv6, k3x3, SE0.25 | 1      | 160      | 9      |
| 6      | MBConv6, k3x3, SE0.25 | 2      | 256      | 15     |
| 7      | Conv1x1&-Pooling      | —      | 1280     | 1      |
| Output | Fully-connected layer | —      | 20       | 1      |

#### 3.2 视觉识别模型的训练

模型的超参数影响模型的训练速度与效果, 为此针对

模型超参数的设置进行了实验分析。本实验对 3 个超参数进行了调试, 这 3 个超参数分别是学习率 (learning rate, LR)、学习率衰减率 (learning rate decay rate, LRDR) 和训练迭代次数 (Epochs)。为提升学习速度, 同时避免迭代次数过少, 实验中将 batch size 设置为 16。

##### 1) 学习率

学习率, 即每次梯度下降时参数随梯度方向下降的比例。在深度学习的模型中学习率是所有超参数中影响模型训练最重要的一个。为了固定其它参数, 则将 Epochs 先设定为 15, 学习率衰减率设为 1, 即全局学习率保持不变, 小节 3) 将会进行解释。

##### 2) Epochs

在深度学习的模型中 Epochs 是控制训练次数的一个超参数, Epochs 即遍历整个数据集的次数, 遍历一次数据集即使用整个数据集对模型进行一次完整训练。实验中每训练一个 Epoch, 都将训练集按照 203 个 batch 进行一次完整训练。当 Epochs 设置过大时训练耗时长, 有可能会让模型产生过拟合 (over-fitting) 的现象, 也会让调节其他超参数更加麻烦; 而当 Epochs 过小的时候, 会由于模型未完成训练而导致欠拟合 (under-fitting), 从而影响测试集的准确率 (accuracy rate, AR)。由于学习率衰减方程和 Epochs 相关, 所以需要先调节出合适的 Epochs 再对学习率衰减进行调节, 同样设置学习率衰减率为 1。

##### 3) 学习率衰减率

在模型的训练的初始阶段, 梯度方向一般和最佳拟合的参数方向一致性较高, 所以在训练开始阶段使用较大的学习率有利于加快学习速度, 而在训练的后期阶段, 梯度方向和最佳拟合的参数方向有可能会有巨大的差距, 所以使用较小的学习率会使模型更加容易拟合。因此, 本实验选择使用学习率 cosine 衰减, 其衰减方程可以表示为:

$$LR = init\_LR \times f_{LR}(ep) \tag{2}$$

$$f_{LR}(ep) = \frac{1}{2} \left( 1 + \cos \left( \frac{ep}{eps} - \pi \right) \right) (1 - LRDR) + LRDR \tag{3}$$

其中, init\_LR 初始学习率即模型的学习率, 本实验设置为 0.01, ep 为即当前训练的 Epoch 序号, eps 为总 Epoch 数量, LRDR 为学习率衰减率。

图 10 是 LR=0.01、LRDR=0.01 时的学习率变化曲线, 其图像为一个 cosine 型曲线。当 Epoch=0 时, 学习率当前值等于初始学习率即 0.01, 当 Epoch=99 时, 学习率当前值等于学习率乘以学习率衰减率。

##### 4) 模型优化器

本实验所使用的优化器 (optimizer) 为 SGD 优化器其中动量值 (momentum) 设置为 0.9, 动量梯度下降所使用的思想是在每次梯度下降时使用的梯度都会考虑到历史梯度, 其迭代公式为:

$$v_i = v_{i-1} \times momentum + g_i \tag{4}$$

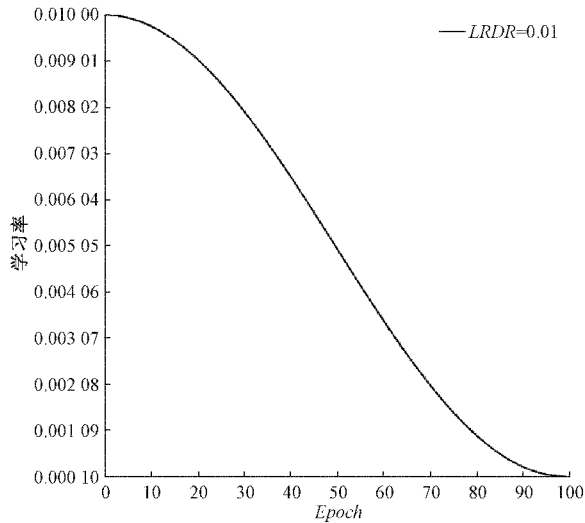


图 10 学习率衰减方程曲线

## 4 实验结果及分析

### 4.1 实验结果评价标准定义

对于样本均衡的分类任务实验结果最准确的评价指标为准确率 (accuracy rate, AR) 和交叉熵损失函数 (cross entropy loss, CE) 其定义如下。

1) 准确率

分类模型识别总结果中正确识别的部分所占比例, 即:

$$AR = \frac{T}{T+F} \quad (5)$$

其中,  $T$  (true) 代表正确分类样本的个数,  $F$  (false) 代表错误分类样本的个数。准确率越高代表本实验的分类效果越好, 其完全正确分类时的理想值为 1(100%)。

2) 交叉熵损失函数<sup>[17-18]</sup>

在经过分类网络之后, 会得到每一个样本的预测结果的分数值 (logits), 在实验中为一个  $20 \times m$  的矩阵  $\mathbf{a}$  ( $m$  为样本数), 经过 softmax 函数之后会得到一个同形状的概率矩阵  $\mathbf{S}$ ,  $j$  代表矩阵的第  $j$  行,  $N$  是分类数即 20, 最后  $\mathbf{S}$  作为交叉熵损失函数的输入, 输出为  $\mathbf{L}$ , 交叉熵损失能直接反应预测的结果和实际结果之间的差距。

softmax 函数定义:

$$S_j = \frac{e^{a_j}}{\sum_{k=0}^{N-1} e^{a_k}} \quad (6)$$

交叉熵损失函数定义:

$$L = - \sum_{j=0}^{N-1} y_j \log S_j \quad (7)$$

### 4.2 调节学习率

通过设置学习率分别为 0.1、0.01、0.001 进行训练, 可以看到, 如图 11 所示, 3 个学习率在训练过程中损失均快速下降, 但相对于学习率为 0.1 和 0.001, 学习率为 0.01

时, 训练损失下降速度是最快的即模型的学习速度是最快的, 所以本实验训练模型选择学习率等于 0.01。

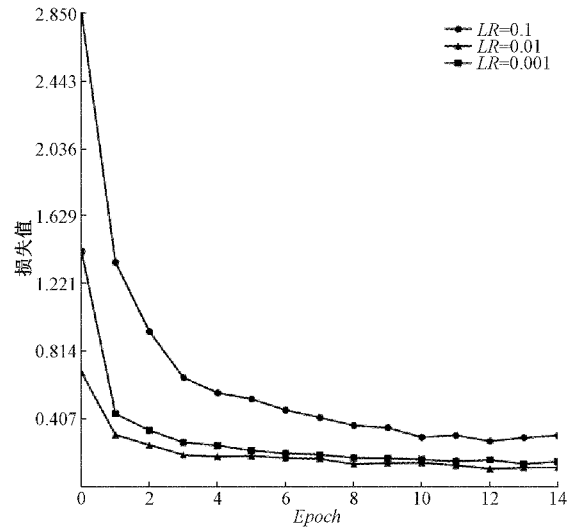


图 11 不同学习率下训练损失对比

### 4.3 调节学习率衰减率

从图 12 来看, 在前 50 次 Epochs 训练中, 不同的学习率衰减率下, 训练模型的验证集损失均快速下降至  $10^{-5}$  左右。在 50 Epoch 之后的迭代中,  $LRDR=1$  时, 损失没有进一步下降, 而  $LRDR=0.1$  和  $LRDR=0.01$  情况下, 模型的损失仍进一步缓缓下降。可见,  $LRDR=0.1$  和  $LRDR=0.01$  能够进一步降低模型验证集损失, 提升模型性能, 而在训练后期  $LRDR=0.01$  时损失能够降到更低的水平所以实验选择设置学习率衰减率为 0.01。

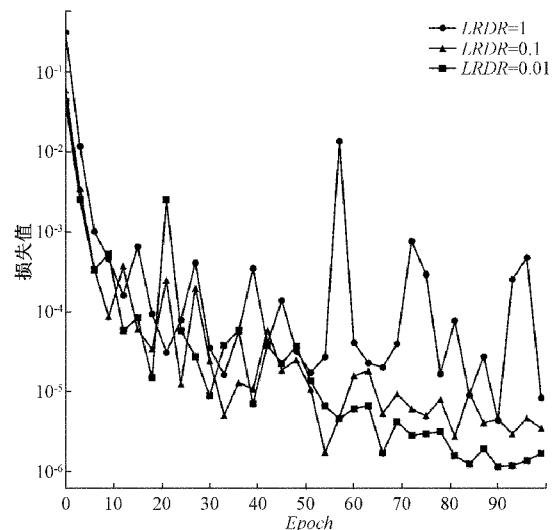


图 12 不同学习率衰减率下验证集损失对比

### 4.4 Epochs 参数调整

在实验过程中, 先设置 Epochs 为 100 可以看到, 如图 12 所示, 当 Epochs=15 的时候损失还在继续下降, 模型明显未完成训练; 当 Epochs=50 的时候模型接近拟合但损失还

能继续下降,而当 Epochs 在 60~90 的时候损失在一定范围内波动,但整体仍然处于下降趋势,而当 Epochs 大于 90,损失没有明显变化,说明模型收敛,继续训练计算代价大,但实际意义不大。因此,本实验的 Epochs 保持设置为 100。

4.5 不同模型对比

为比较 EfficientNetV2 在实验中作为视觉识别模型的有效性,故实验设置 ResNet<sup>[19]</sup> 和 GoogLeNet<sup>[20]</sup> 作为对照组模型进行比较。ResNet 是何凯明在 CVPR2016 发表的获得最佳论文奖的一个分类算法,在使用深度学习解决分类问题中(如医学、工程上)得到了广泛应用。GoogLeNet 则是基于 Inception 模块的神经网络模型,它在 2014 年 ImageNet 竞赛夺得了冠军。

在训练阶段,通过将训练的每个 Epoch 训练集与验证集的表现结果及这个 Epoch 训练后的权重保存至本地,并在测试集评估阶段将训练的每个 Epoch 的权重加载进入模型测试集进行测试,最终获得模型训练损失、验证集损失、测试集损失以及测试集准确率。

如图 13 及 14 所示,在训练条件以及超参数均是最优的情况下,ResNet101 在训练过程的第 0 到第 30 个 Epoch 阶段,测试集损失迅速下降,准确率也迅速提升,最终损失稳定在了 0.09 附近,准确率稳定在 98.5% 附近;GoogLeNet 在训练过程的第 0 到第 68 个 Epoch 阶段,测试集损失和准确率均上下大幅度波动,之后的训练中损失维持稳定在 10<sup>-5</sup> 附近,准确率接近 100%;而 EfficientNetV2 在第 4 个 Epoch 之后测试集准确率一直是 100%,测试集损失在第 2 个 Epoch 就已经能够降到 0.01 以下并持续降低,最终稳定在 3×10<sup>-6</sup> 附近,全程表现结果不论是收敛速度,还是最终的测试集准确率与测试集损失,EfficientNetV2 显著优于 ResNet101 与 GoogLeNet。

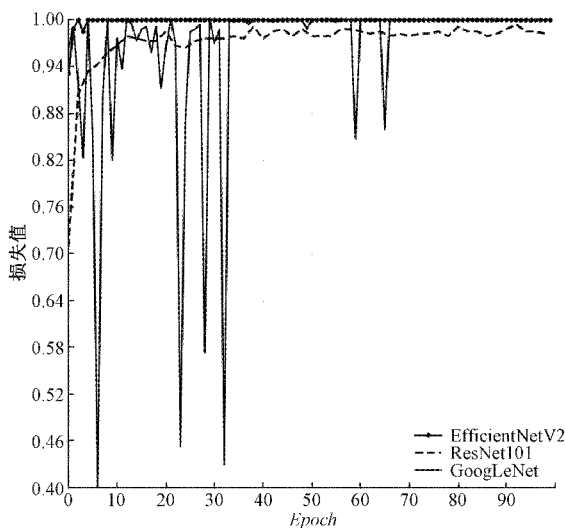


图 13 不同模型测试准确率对比

如表 3 所示, EfficientNetV2 训练 100 个 Epoch 的平均准确率高达 99.9%, 高于 ResNet 和 GoogLeNet; 平均损

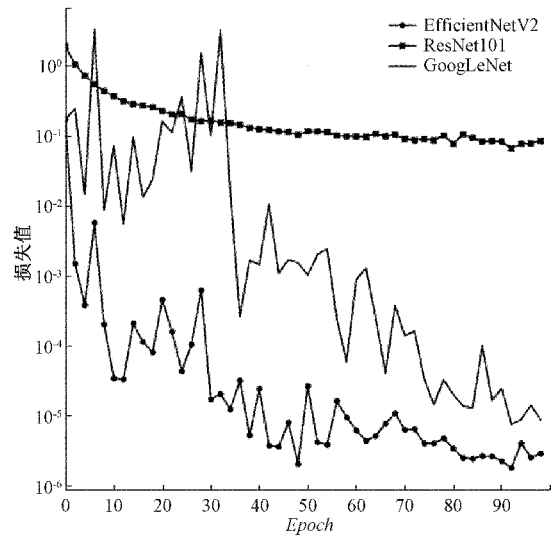


图 14 不同模型测试损失对比

失为 2.61×10<sup>-3</sup>, 最优损失值为 1.70×10<sup>-6</sup>, 远远低于 ResNet 和 GoogLeNet, 其表现效果远远强过 ResNet101 和 GoogLeNet。

表 3 100 个 Epoch 下不同网络效果对比

| 模型             | 平均<br>准确率/<br>% | 最优<br>准确率/<br>% | 平均<br>损失值             | 最优<br>损失值             |
|----------------|-----------------|-----------------|-----------------------|-----------------------|
| EfficientNetV2 | 99.9            | 100             | 2.61×10 <sup>-3</sup> | 1.70×10 <sup>-6</sup> |
| ResNet         | 97.3            | 99.3            | 2.03×10 <sup>-1</sup> | 6.64×10 <sup>-2</sup> |
| GoogLeNet      | 96.2            | 100             | 1.6×10 <sup>-1</sup>  | 3.75×10 <sup>-6</sup> |

5 结 论

本文提出了一种基于双激光照射的扩增机器视觉视野深度的方法,搭建了基于特征扩增 EfficientNetV2 的视觉识别系统,设计了自动旋转移移动采集装置有效地提高了数据规模,训练中引入了学习率 cosine 衰减,并对各种超参数进行调优,有效地提高了视觉识别检测系统的性能。实验选择 20 种表面视觉性状难以区分的液体作为对象,测试准确率能够达到 100%。本文解决了液体视觉检测中特征少的问题,对机器视觉在液体检测领域的发展具有重要意义。

参考文献

[1] 冯国红, 朱玉杰, 杨慧敏, 等. 基于 PCA 的“光纤液滴指纹图”识别液体研究[J]. 现代电子技术, 2019, 42(15): 47-50.

[2] 胡紫薇, 项安, 陈奕杰. 基于双能 X 射线的液体危险品识别分类方法[J]. 无损检测, 2020, 42(9): 63-68.

[3] 王宇石, 杨立瑞, 李保磊. 基于 X 射线图像的容器中液体物质的识别[J]. 中国体视学与图像分析, 2011, 16(1): 33-37.

- [4] 张飞, 耿红琴, 陈萍. 采用虚拟仪器的近红外无损液体识别电子舌设计[J]. 激光杂志, 2015, 36(1): 74-77.
- [5] YAN J, JIN M, XU Z, et al. Recognition of suspension liquid based on speckle patterns using deep learning[J]. IEEE Photonics Journal, 2020, DOI: 10.1109/JPHOT.2020.3044912.
- [6] CAO P, ZHU Y, ZHAO W, et al. Chromaticity measurement based on the image method and its application in water quality detection[J]. Water, 2019, 11(11): 2339.
- [7] GRANT-JACOB J A, PRAEGER M, LOXHAM M, et al. Determination of size of urban particulates from occluded scattering patterns using deep learning and data augmentation[J]. Environmental Research Communications, 2021, 3(2): 025003.
- [8] 权建军, 穆海超, 杨博. 动态液体颜色变化机器视觉识别传感系统设计[J]. 兰州石化职业技术学院学报, 2021, 21(2): 11-14.
- [9] 刘一鸿. 基于机器视觉技术的白酒杂质检测系统研究[D]. 西安: 西安工业大学, 2021.
- [10] TAN M, LE Q V. EfficientNetV2: Smaller models and faster training[C]. Proceedings of the 38th International Conference on Machine Learning, PMLR, 2021, 139: 10096-10106.
- [11] 李航宇, 王楠楠, 朱明瑞, 等. 神经网络结构搜索的研究进展综述[J]. 软件学报, 2022, 33(1): 129-149.
- [12] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, 2020.
- [13] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
- [14] HUANG G, SUN Y, LIU Z, et al. Deep networks with stochastic depth[C]. European Conference on Computer Vision. Springer, Cham, 2016: 646-661.
- [15] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
- [16] 蓝金辉, 王迪, 申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 167-182.
- [17] 石晨宇, 周春, 靳鸿, 等. 基于卷积神经网络的农作物病害识别研究[J]. 国外电子测量技术, 2021, 40(9): 93-99.
- [18] 刘正琼, 丁力, 凌琳, 等. 基于字符编码与卷积神经网络的汉字识别[J]. 电子测量与仪器学报, 2020, 34(2): 143-149.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [20] SZEGEDY C, WEI L, JIA Y, et al. Going deeper with convolutions[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

#### 作者简介

钟扬, 硕士研究生, 主要研究方向为机器视觉。

E-mail: 769012871@qq.com

吴黎明, 教授, 主要研究方向为智能仪器。

E-mail: jkyjs@gdut.edu.cn

温腾腾, 博士后, 主要研究方向为智能感知。

E-mail: wentt@gdut.edu.cn

伍冠楚, 硕士研究生, 主要研究方向为机器视觉。

E-mail: 549713043@qq.com

王桂棠, 教授, 主要研究方向为智能测控。

E-mail: wanggt@gdut.edu