

基于 U-P-Net 的手机玻璃屏幕缺陷分割

李 墨¹, 陈志豪², 张 颢³

(1. 广东工业大学 计算机学院, 广州 510006; 2. 广东工业大学 自动化学院, 广州 510006;
3. 华南理工大学 电力学院, 广州 510641)

摘要: 随着科技的发展及电子设备的普及, 玻璃屏幕质量成为电子设备和其他产品的重要考虑因素; 而玻璃外观缺陷检测是玻璃质量检测中最重要的一环, 这也是保证产出高品质、高性能的玻璃产品的关键环节; 目前玻璃表面缺陷检测方法存在无目标训练图像资源消耗、检测精度较低、复杂特征信息难以提取等问题; 因此, 为了解决上述问题, 提出了一种基于 U-pyramid pooling module-Net (U-P-Net) 的手机玻璃屏幕缺陷分割模型; 采用超像素预处理, 有效地降低了原始图像的复杂度; 采用 ResNet50 作为分类网络, 减少无目标训练图像造成的资源浪费, 提高训练效率; U-P-Net 被提出, 有效地聚合了不同区域的上下文信息, 提高了获取全局信息的能力; 实验结果表明, 所设计的基于 U-P-Net 玻璃缺陷分割算法分割精度明显优于其它传统卷积神经网络分割方法, 证明了该框架在移动屏幕数据集上的有效性。

关键词: 缺陷检测; 超像素预处理; ResNet50; U-P-Net; 金字塔池化; 手机玻璃屏幕

Defect Segmentation of Mobile Phone Screen Based on U-P-Net

LI Mo¹, CHEN Zhihao², ZHANG Xie³

(1. School of Computer Science, Guangdong University of Technology, Guangzhou 510006, China;
2. School of Computer Automation, Guangdong University of Technology, Guangzhou 510006, China;
3. School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, China)

Abstract: With the development of science and technology and the popularity of electronic equipment, glass screen quality has become an important consideration for electronic equipment and other products; The detection of glass appearance defects is the most important link in glass quality detection, which is also the key link to ensure the production of high quality and high-performance glass products; At present, there are some problems in the detection methods of glass surface defects, such as resource consumption of target-free training image, low detection accuracy and difficult extraction of complex feature information. Therefore, to solve the above problems, a defect segmentation model of mobile phone glass screens based on U-pyramid pooling module-Net (U-P-Net) is proposed. Superpixel preprocessing is used to reduce the complexity of the original image effectively. ResNet50 was used as a classification network to reduce the waste of resources caused by training images without targets and improve training efficiency. U-P-Net is proposed, which aggregates the context information of different regions effectively and improves the ability to obtain global information. Experimental results show that the proposed U-P-Net glass defect segmentation algorithm is significantly superior to other traditional convolutional neural network segmentation methods, which proves the effectiveness of the framework on mobile screen data sets.

Keywords: defect detection; superpixel pre-processing; ResNet50; pyramid pooling; mobile phone glass screen

0 引言

随着科技的不断进步及工业的快速发展, 玻璃制品已经成为众多行业中不可或缺的重要组成部分。玻璃作为一种需求量日益增加的特殊材料, 细微缺陷就会影响其机械特性和光学性质, 从而降低甚至破坏玻璃产品的使用价值。因此, 玻璃外观质量检测是玻璃质量检测中最重要的一环, 这也是保证产出高品质、高性能的玻璃产品的关键环节。传统的玻璃面板缺陷检测方法主要依赖于高级技工的人工检验。随着玻璃生产制造需求逐渐增大, 传统的人工检验方法无法满足高速生产带来的庞大的质量检测需求, 因此

发展高质量的自动化玻璃缺陷检测系统变得尤为重要。随着人工智能技术的发展, 许多大型企业在玻璃面板缺陷检测中引入机器视觉方法代替人工检测, 以提高企业的产品质量和生产效率。但机器视觉的应用也存在一些弊端, 它不仅需要针对特定的应用场景定制, 无法适应一系列复杂动态的工业制造场景, 而且缺乏自主学习和技能迁移能力^[1-9]。

近年来, 在深度学习算法中, 基于卷积神经网络的语义分割方法在缺陷检测方面得到了广泛的应用^[10-17]。该方法可以准确提取复杂背景中的特征, 快速分割人体肉眼

收稿日期: 2023-03-02; 修回日期: 2023-03-14。

基金项目: 国家重点研发计划项目(2020AAA0108304); 国家自然科学基金项目(62073088, U1911401)。

作者简介: 李 墨(1999-), 女, 湖南岳阳人, 硕士研究生, 主要从事深度学习目标检测、图像处理方向的研究。

引用格式: 李 墨, 陈志豪, 张 颢. 基于 U-P-Net 的手机玻璃屏幕缺陷分割[J]. 计算机测量与控制, 2023, 31(8): 231-237.

难以识别的不明显缺陷,从而实现玻璃面板的有效缺陷检测。2015年,Ronneberger等^[18]提出了一种名为U-Net的网络,该网络基于语义分割技术,可以实现对生物医学图像进行像素级别的分类。然而此网络模型的学习参数较多,训练难度较大,且需要进行大量的图像标注工作;参考文献19利用了卷积神经网络中的VGG Net模型对手机屏幕缺陷进行了检测,然而该研究发现,在4种主要缺陷的检测中,漏检率达到了6%~9%,最高的检测率仅有94%,这种检测精度无法满足相关领域的高要求^[19];针对现有的语义分割方法在面对不规则数据时存在的计算复杂度高、语义特征信息提取不充分问题。提出了一种结合超像素技术与残差分类网络的预处理方法,同时,将金字塔池化模块PPM嵌入到U-Net模型中,构成了一种全流程的新的语义分割方法。

1 关键技术

目前的语义分割模型主要存在以下几个问题:

- 1) 缺陷样本图像的高度复杂性和可变性使得提取特征信息具有挑战,导致分割性能差。
- 2) 由于缺陷样本图像的特殊性非目标训练图像占用了训练资源,导致训练速度变慢,损失函数收敛速度变慢。
- 3) 多变场景时获取全局信息不充分,造成语义分割时出现不匹配关系、混淆类别和不显著类别的缺陷误判,未判和错判等问题。

本文模型基于经典U-Net模型,对输入图像进行数据预处理,并使用PPM对U-Net网络特征处理层进行融合,有效降低了模型对数据集的依赖,降低模型的训练复杂度、提高了分割精度和速率。其总体模型框架如图1所示。

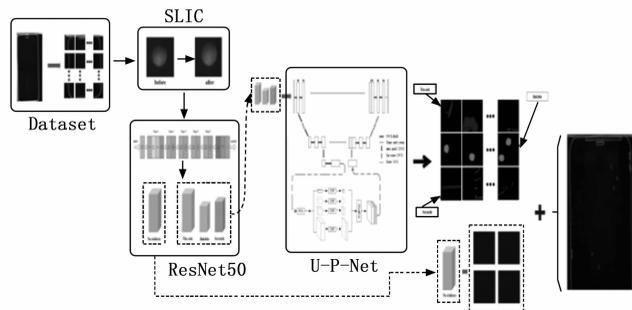


图1 总体模型框架图

本文主要关键技术如下:

- 1) 由于图像特征信息的复杂性和可变性使得特征信息的提取具有挑战性,导致分割性能较差。因此,本文提出采用对输入图像进行超像素预处理。超像素^[20-22]是一个5维特征向量,它将彩色图像转换为CIELAB颜色空间和XY坐标。此外,对5维特征向量构建距离测量准则,对图像像素进行局部聚类。且SLIC^[20]是基于纹理、颜色、亮度等相似的相邻像素快速生成紧凑干净的超像素块,通过构造一个简单的线性迭代聚类图像预处理模块,解决了图像特征信息复杂多变的问题,使邻域特征表示更加容易。它不仅可以降低图像的维数,还可以捕获冗余信息,消除像素异常点。

不仅可以降低图像的维数,还可以捕获冗余信息,消除像素异常点。

- 2) 由于输入图像的特殊性,在语义分割前对缺陷图像进行分类,能在一定程度上降低后续图像语义分割的复杂度,提升网络模型的训练速度及分割准确率。该方法使用带有残差结构ResNet50^[23-26]网络对缺陷图像进行初步分类。其独特的残差结构创建了不同网络层之间的快速连接(shortcut connection)。它解决了深层网络中梯度弥散和精度下降(训练集)的问题,使网络能够越来越深,既保证了精度,又控制了速度。

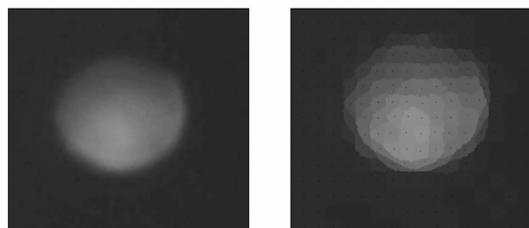
- 3) 全卷积网络(FCN)的缺点是需要一种合适的策略来利用全局场景分类信息的特征。为了解决上述问题,提出了U-pyramid pooling module-Net(U-P-Net)。该方法引入金字塔池化模块PPM^[27-28],在4个不同的粗细尺度上进行特征融合,利用多样化的卷积尺寸,提取到的特征更具多样性,且能够聚合不同区域的上下文信息,进一步提高获取全局信息的能力,避免关系错误匹配、类别混淆以及细小对象的类别被忽略问题。

2 模型介绍

该框架由3个模块组成:SLIC超像素预处理模块^[22]、分类网络模块和U-P-Net模块。每个模块将在以下部分中详细描述。

2.1 超像素预处理模块

手机玻璃屏幕数据集中样本图像的高复杂性和高分辨率使得提取特征信息具有挑战性,导致分割性能较差。考虑到图像中目标特征的复杂性,采用超像素算法进行图像预处理。SLIC^[22]通过构造一个简单的线性迭代聚类图像预处理模块,解决了图像特征信息复杂多变的问题。SLIC^[22]基于纹理、颜色、亮度等相似的相邻像素快速生成紧凑干净的超像素块,使邻域特征表示更加容易。SLIC^[20]预处理效果如图2所示。



(a)原始气泡缺陷图像

(b)SLIC预处理后气泡缺陷图像

图2 SLIC预处理效果图

SLIC具体步骤如下:

- 1) 初始化(撒种子):根据图像大小及其他参数确定超像素数目,在图像内均匀分配种子。
- 2) 聚类中心选择:在聚类中心内计算所有像素点的梯度值,并选择梯度最小的像素点作为新的聚类中心。
- 3) 分配像素到超像素:每个像素必须与之位置重叠的最近聚类中心之间相关联,将其分配到间距最近的超像素。

4) 距离的度量: 包括像素之间的颜色距离和空间距离, 其度量方法如下:

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \quad (1)$$

$$d_{xy} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (2)$$

$$D = \sqrt{\left(\frac{d_{Lab}}{N_c}\right)^2 + \left(\frac{d_{xy}}{N_s}\right)^2} \quad (3)$$

式中, d_{Lab} 为颜色距离, d_{xy} 为空间距离, $N_s = \sqrt{N/K}$ 每个集群均适用, 是类内最大空间距离。 N_c 是最大的颜色距离, 随着图像和集群的不同而不同, 可使用固定常数 m 来调节 d_{Lab} 与 d_{xy} 的影响权重, 取值范围为 $[1 \sim 40]$, 默认取值为 10。 D 为最终的距离度量, 计算方法如下所示:

$$D = \sqrt{\left(\frac{d_{Lab}}{m}\right)^2 + \left(\frac{d_{xy}}{N_s}\right)^2} \quad (4)$$

5) 迭代过程: 重复执行 3) 和 4), 直到超像素中心位置不再发生变化或达到最大迭代次数。

6) 聚类优化: 解决经过迭代后存在的与聚类中心不属于同一连通域的孤立像素点。具体步骤为: 利用连通算法将相邻的超像素合并成一个超像素, 如果合并后的超像素与某个邻居超像素的距离小于一定阈值, 则将其合并。

2.2 分类网络模块

手机玻璃屏幕数据集中存在一些非目标训练图像, 这些图像会占用训练资源, 导致训练速度变慢, 并且损失函数的收敛速度也变慢。这种情况会严重影响训练效率和准确性, 因此需要采取相应措施来解决非目标训练图像对训练过程的影响, 提高训练效率和准确性。本文选用了一种名为 ResNet50^[23] 的残差网络作为手机玻璃屏幕图像的分类网络。 ResNet50 是一种深度残差网络, 具有易优化、计算负担小等优点, 在计算机视觉领域得到了广泛的应用。与传统的卷积神经网络相比, 残差块 (residual block) 的引入是 ResNet50 解决梯度消失和退化问题的重要方法之一, 它有效地提高了网络的性能。

具体来说, 残差块通过引入跳跃连接来绕过了某些卷积层和激活函数层, 线性层拟合残差映射, 使得神经网络的信息流能够更加顺畅地传递。这种结构的引入使得 ResNet50 网络可以更加深层次地学习特征, 达到充分训练底层网络的效果, 同时减轻了梯度消失和过拟合的问题。 ResNet50 网络结构如图 4 所示, 由 49 个卷积层和 1 个全连接层组成。其中, 第二至第五个阶段中的 ID BLOCK x2 是指包含两个不改变维度的残差块, 而 CONV BLOCK 则是指包含添加维度的残差块。每个残差块均由 3 个卷积层组成, 因此整个网络包含 49 个卷积层。

2.3 U-pyramid pooling- Net 模块

手机玻璃屏幕在生产过程中容易受到环境的影响, 导致产生气泡、划痕、锡灰等各种形状和大小的缺陷。这些缺陷的语义分割存在关系不匹配、混淆和对象类别小等问题, 给模型的准确分割造成了挑战。为了解决上述问题, 提出了 U-P-Net 模块。该模块可以有效地处理这些异形的

缺陷, 提高语义分割的准确性和可靠性。

U-P-Net 的结构如图 5 所示。通过在 U-Net 网络的主干特征提取部分和增强特征提取部分之间引入了金字塔池化模块, 以进一步提高网络的性能。该模块可以将不同大小的特征图进行池化, 从而捕获更多的上下文信息, 提高网络的感受野和特征表达能力, 避免语义分割问题中的误分类、错分类、未分类等问题。

PPM^[27] 的结构如图 3 所示, 是一种多尺度特征融合的方法, 其基本原理是在原始特征图上进行 4 个不同粗细尺度的池化操作, 然后将不同尺度的特征融合在一起, 其中最粗尺度产生单格输出; 加细尺度产生多格输出。金字塔池化会对输入图像进行多次下采样, 从而得到不同尺度的特征图, 然后低维特征图通过双线性插值进行上采样, 以获得相同大小的特征图。最后, 将这些特征图 (包括原始特征图) 在通道维度上进行拼接, 生成一个综合了多种尺度信息的复合特征图, 从而达到兼顾全局语义信息与局部细节信息的目的。

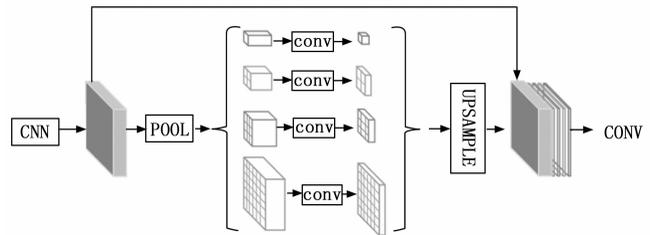


图 3 PPM 金字塔池化模块结构图

U-P-Net 具体步骤如下:

1) 输入图像进入主干特征提取网络, 由两个 3×3 的卷积层 (ReLU) 再加上一个 2×2 的 maxpooling 层组成一个下采样的模块, 一共经过 4 次这样的操作, 得到初步有效特征层 feature map;

2) 对骨干提取的 feature map (channel = N) 做池化得到特征金字塔;

3) 然后通过 1×1 深度卷积降通道分别得到 1×1 , 2×2 , 4×4 , 6×6 的 channel = $1/N$ 的特征图;

4) 对特征图进行双线性插值填充上采样到原 Feature map 尺寸;

5) 与 Feature map 进行通道拼接 (也叫级联) 得到 channel 数增加 1 倍的特征图;

6) 再利用 1×1 卷积核将上述拼接后的特征图深度卷积降通道得到与输入特征图 Feature 通道数一致的结果得到最终的金字塔池化全局特征;

进入加强特征提取部分, 由一层反卷积 + 特征拼接 concat + 两个 3×3 的卷积层 (ReLU) 反复构成, 一共经过 4 次这样的操作, 对得到的金字塔池化全局特征进行上采样, 并且进行特征融合, 最后接一层 1×1 卷积, 降维处理即将通道数降低至特定的数量, 获得一个最终的, 融合了所有特征的有效特征层的目标图。

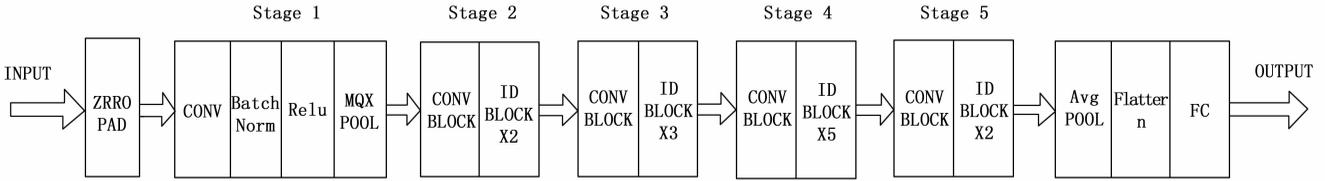


图 4 ResNet50 网络结构图

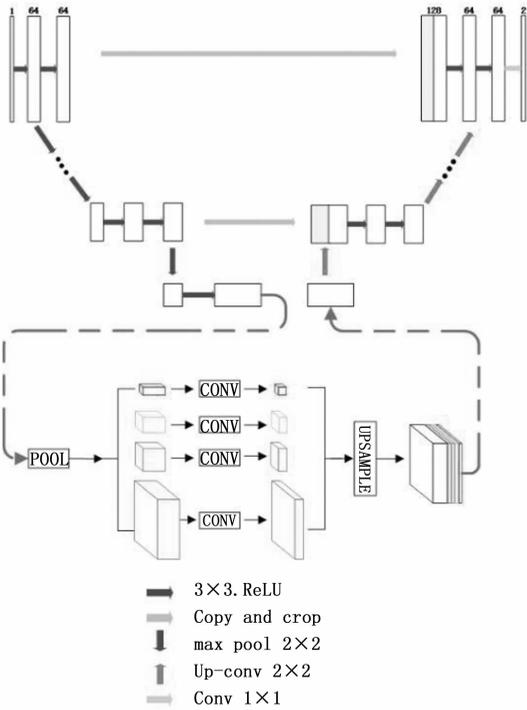


图 5 U-P-Net 模型结构图

2.4 损失函数

损失函数是对框架的预测值与实际值不一致的计算。它是非负的实值函数。损失函数越小，框架的训练性能越好，鲁棒性越好。损失函数是训练结果连续接近实际值时必须使用的函数。其中，损失函数是神经网络通过梯度下降等方法调整权重参数。本文框架中使用的损失函数为加权交叉熵损失函数：

$$E = \sum_{x \in \Omega} \omega(x) \log(P_{l_i}(x)) \quad (5)$$

其中： $\Omega \rightarrow \{1, \dots, K\}$ 为像素点的标签值， $p_{l_i}(x)$ 为 softmax 损失函数。 ω 为 $\in R$ 为像素点的权值，为图像中靠近边界点的像素赋予更高的权值。

$$p_{l_i}(x) = \exp(a_{l_i}(x)) / (\sum_{l'=1}^k \exp(a_{l'}(x))) \quad (6)$$

其中： $p_{l_i}(x)$ 为 softmax 函数。 $a_{l_i}(x)$ 是第 k 个特征通道（类别的数量等于特征通道的数量，总共 k 个）。

$$\omega(x) = \omega_c(x) + \omega_0 \cdot \exp\left(-\frac{(d_1(x) + d_2(x))^2}{2\sigma^2}\right) \quad (7)$$

其中： ω_c 为 $\in R$ 是平衡类别比的权值， d_1 为 $\in R$ 是像素点到其最近像素点的距离， d_2 为 $\in R$ 是像素点到距离其第二

近像素点的距离。 ω_0 和 σ 为常数，实验中 $\omega_0 \approx 10$ ， $\sigma \approx 5$ 。

3 实验

3.1 实验数据

本实验中使用的手机玻璃屏幕数据集是在实验室液晶显示器 (LCD) 屏幕缺陷检测设备上采集的手机玻璃屏幕图像。数据集包含 3 种类型的缺陷：划痕 (scratch)、气泡 (bubble) 和锡灰 (tin_ash)。图 6 显示了 3 种手机玻璃屏幕的表面缺陷样例。

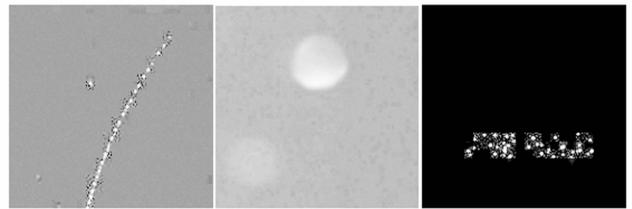


图 6 手机玻璃屏幕缺陷划痕、气泡和锡灰样例图

由于原始手机玻璃屏幕数据集的图像分辨率为 6400×6400 ，分辨率过高，难以分割缺陷且分割效果较差。因此，对原始图像进行图像分割和数据增强，得到分辨率为 640×640 的图像。并对其进行数据增强和数据扩充，其中包括增加噪声、亮度调整、图像模糊和图片旋转等方法，以解决样本分辨率过高以及样本不均衡所导致的分割效果较差和模型泛化能力低的问题。

本实验设置手机玻璃屏幕数据集来验证本文框架的性能。手机玻璃屏幕数据集的训练集、验证集和测试集的数量分布如表 1 所示。该数据集是一个用于手机玻璃屏幕缺陷语义分割的图像数据集，其中包含了各种常见的屏幕缺陷，如气泡、划痕、锡灰。数据集的训练集共有 600 张图像，用于训练模型的参数。验证集共 100 张图像，用于选择最佳的模型超参数。测试集共有 200 张图像，用于测试模型的性能表现。在训练集中，各种缺陷的数量分布如表二所示。气泡缺陷的数据图像共 206 张，划痕缺陷的数据图像共 167 张，锡灰缺陷的数据图像共 227 张。

表 1 手机玻璃屏幕数据集的数据分布

Training set	Validation set	Testing set	Dataset
600	100	200	900

表 2 手机玻璃屏幕训练集缺陷分布

Bubble	Scratch	Tin-ash
206	167	227

3.2 评估指标

在实际的语义分割任务中, 语义分割框架的预测性能评价指标为交集除以并集 (IoU)、像素精度 (PA)、平均交集过并 (mIoU) 和平均像素精度 (mPA)。在性能评估计算公式中, 真正 TP (true positive) 是指对某一类缺陷, 语义分割框架正确分割出的缺陷数量。假阳性 (FP) 是指语义分割框架对这种缺陷错误所分割出的缺陷数量。假阴性 (FN) 是没有被语义分割框架分割的缺陷的数量, 对于这类缺陷是正确的。真负 (TN) 是没有被语义分割框架分割且不属于这种类型的缺陷的数量, 通常为 0。

1) IoU 是像素点的真值与预测值的交集除以像素点的真值与预测值的拼接。计算公式如下:

$$IoU = \frac{\sum_{i=0}^k p_{ij}}{\sum_{i=0}^k (\sum_{i=0}^k p_{ii} + \sum_{j=0}^k p_{ji} - p_{ii})} = \frac{\sum_{i=0}^k TP}{\sum_{i=0}^k FN + FP + TP} \quad (8)$$

2) PA 用于计算正确分类的像素数与像素总数的比值。计算公式如下:

$$PA = \frac{\sum_{i=0}^k P_{ii}}{\sum_{i=0}^k \sum_{j=0}^k P_{ij}} \quad (9)$$

3) mIoU 是数据集中各类相交比的平均值。计算公式如下:

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ij}}{\sum_{j=0}^k p_{ii} + \sum_{j=0}^k p_{ji} - p_{ii}} = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN + FP + TP} \quad (10)$$

4) mPA 分别是计算每一类正确分类像素的比例。计算公式如下:

$$mPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \quad (11)$$

其中: i 为真实值。 j 表示预测值。 p_{ij} 表示 i 作为 j 的预测, k 是类别的数量。

3.3 训练细节

该液晶屏缺陷检测设备由运动控制模块、图像采集模块、光源控制模块、信号产生模块组成。运动控制模块采用三菱 PLC 和伺服驱动进行运动控制。图像采集模块采用海康威视 CH 10 GbE 工业面阵相机。光源控制模块采用双面低角度 LED 光源作为检测环节的除尘灯。信号产生模块采用信号发生器模拟手机主板给予手机玻璃面板的电信号, 实现面板的不同颜色变换。其中信号产生模块只选择了一个场景。

为了加快网络的训练速度, 本文采用单个 GPU (NVIDIA GeForce RTX 3 080) 来提高计算效率。使用 Python 实现改进 U-Net^[18]。提出的框架是用 python 进行训练和测试的。整个实验采用开源深度学习框架 PyTorch 实现。初始获取的每张图像大小为 640×640 。优化算法采用 Ad-

am 算法, 发生器和鉴别器学习率均设置为 0.000 1, beta1 为 0.5; 超参数设置为 0.5; 训练步数设置为 3 000 步。在获得足够的数后, 使用自适应矩估计 (Adam) 优化算法进行训练, 动量为 0.9, 权值衰减为 0.000 1, beta1 为 0.9, beta2 为 0.999。学习率为 1×10^{-3} 。训练批数设置为 16, 以避免局部极小值。

3.4 对比实验

在本节中, 我们验证了本文提出的框架在手机玻璃屏幕数据集上的有效性, 并将其性能与其他流行的语义分割框架进行了比较。实验涉及的主流框架为 U-Net^[18]、PSP-Net^[29] 和 DeepLab V3^[30]。

图 7 显示了本文提出的框架在标准手机玻璃屏幕数据集下的收敛曲线。由图 7 可知, 曲线在大约 800 Epoch 收敛。

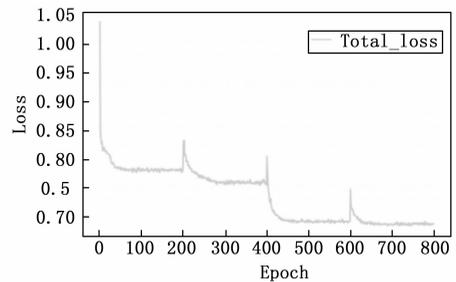


图 7 模型训练中框架的收敛曲线

本文提出的框架是针对手机玻璃屏幕缺陷检测任务设计的, 经过在标准数据集上的训练和测试, 取得了较为优秀的性能表现。各缺陷的 IoU 和总 mIoU 指标如图 8 所示, PA 和总 mPA 指标如图 9 所示。其中, 对于气泡缺陷, IoU 指标为 57.88%, PA 指标为 59.90%; 对于划痕缺陷, IoU 指标为 52.75%, PA 指标为 54.01%; 对于锡灰缺陷, IoU 指标为 82.25%, PA 指标为 87.59%。此外, 本文提出的框架在背景区域的分割上也取得了较高的 IoU 和 PA 指标, 分别为 95.67% 和 98.12%。

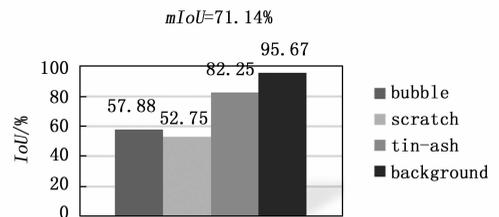


图 8 本框架在标准手机玻璃屏幕数据集中的 MIoU 指标

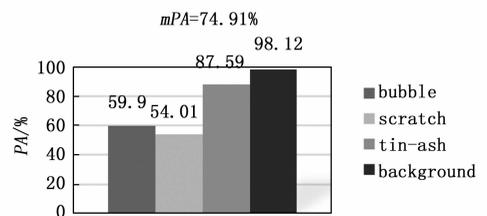


图 9 本框架在标准手机玻璃屏幕数据集中的 mPA 指标

综合各个缺陷类型的性能表现, 本文提出的框架在总 $mIoU$ 和总 mPA 指标上分别达到了 71.14% 和 74.91%, 相比于传统方法和现有的深度学习方法, 具有更高的检测准确率和更强的鲁棒性。

表 3 展示了本文提出的框架与其他主流框架在标准手机玻璃屏幕数据集上各种缺陷语义分割的性能对比。与经典 U-Net^[18] 网络相比, 该框架的 mPA 和 $mIoU$ 性能指标分别提高了 4.89% 和 5.49%。与 PSPNet^[29] 网络相比, 该框架的 mPA 和 $mIoU$ 性能指标分别提高了 46.96% 和 44.25%。与 DeepLabv3^[30] 网络相比, 该框架的 mPA 和 $mIoU$ 性能指标分别提高了 38.87% 和 34.92%。该框架提出的超像素预处理

表 3 本文采用 U-Net、PSPNet 和 DeepLabv3 作为本文提出的框架的比较框架

Number of dataset	Method	Backgr ound/%	Scra tch/%	Bub ble/%	Tin- ash/%	mPA /%	$mIoU$ /%
1 000	U-Net	94.81	62.51	64.85	42.57	71.13	66.19
1 000	PSPNet	91.12	0.00	18.59	0.00	29.06	27.43
1 000	DeepLabv3	92.01	0.23	28.56	26.25	37.15	36.76
1 000	Proposed framework	95.39	49.91	62.11	79.93	76.02	71.68

模块利用超像素预处理和 ResNet50^[23] 残差网络分类, 简化了输入图像的复杂性和冗余度。提高了后续图像缺陷语义分割框架的准确性和鲁棒性。另一方面, 在该框架下提出的 U-P-Net 可以利用多样化的卷积大小提高框架的特征提取能力, 并可以聚合不同区域的上下文信息, 进一步增强获取全局信息的能力, 有效提高分割效率, 优化分割效果。

3.5 消融实验

为了验证本文框架中各个模块的有效性, 整个实验采用了 3 种不同的烧蚀方案。比较了 U-Net、超像素预处理模块、分类网络模块和 U-P-Net 模块。3 种方案在标准移动屏幕数据集上的性能指标如表 4 所示, 其中“+”表示该模块的添加。方案 1 选用经典的 U-Net^[18] 作为整个框架; 方案 2 在 U-Net 语义分割前对图像进行超像素预处理和分类。方案 3 是本文提出的框架。

表 4 使用不同方案在标准手机表面数据集上的性能 %

U-Net	SLIC+ ResNet50	U-P-Net	Backgr round	Scr atch	Bub ble	Tin- ash	mPA	$MIoU$
+	-	-	96.79	63.28	66.58	48.71	74.32	68.84
+	+	-	97.29	55.82	60.23	85.24	76.02	72.89
-	+	+	97.29	56.21	65.01	85.24	80.25	74.96

从表中可以看出, 方案 2 在增加图像数据预处理模块 (SLIC^[22]+ResNet50^[23]) 后, $mIoU$ 和 mPA 的性能较方案 1 分别提高了 4.05% 和 1.7%。

且在加入图像数据预处理模块后模型对锡灰这一类缺陷的分割效果, 大幅度提升。相比 U-Net 网络对锡灰缺陷

的分割, SLIC 和 ResNet50 的加入解决了锡灰缺陷存在的特征信息复杂多变的问题, 使其邻域特征的表示更加容易, 降低了锡灰缺陷的图像复杂度, 提升模型的分割效果。与方案 2 相比, 方案 3 在此基础上引入了 U-P-Net 模块。U-P-Net 模块融合了缺陷图像的深层和浅层特征, 提高了网络获取全局信息的能力, 避免了关系不匹配、类别混淆和小对象类别被忽略等问题。 $mIoU$ 和 mPA 性能较方案 1 分别提高 4.23% 和 2.07%, 较方案 2 分别提高 4.23% 和 2.07%。综上所述, 在分别引入 U-P-Net 模块和图像数据预处理模 (SLIC+ResNet50) 后, 各项指标均有不同程度的提升, 而方案的三项性能明显优于方案一和方案二。

4 结束语

鉴于现有缺陷检测的语义分割算法存在分割不准确、语义信息提取不足、面对不规则数据收敛时间长等问题。提出了一种实用的手机玻璃屏幕缺陷分割框架。在数据预处理方面, 结合超像素技术和残差分类网络的数据预分类方法可以有效降低模型对数据集的依赖, 降低后续图像处理的计算复杂度。U-P-Net 的合理提出, 使所提出的模型能够有效地聚合不同区域的上下文信息, 提高模型获取全局信息的能力, 避免了语义分割中的误分类和漏分类问题。框架在缺乏足够数据集的情况下显示出良好的性能。它在分割效率和精度方面优于许多主流模型, 可以应用于手机玻璃屏幕表面生产。接下来的步骤是如何在分割中更好地结合超像素, 如何在单张图像中分割多个缺陷, 以及如何将其应用于不同材料的图像。

参考文献:

- [1] BODNAROVA A, BENNAMOUN M, LATHAM S. Optimal gabor filters for textile flaw detection [J]. Pattern Recognition, 2002, 35 (12): 2973 - 2991.
- [2] WILDES R P. A machine-vision system for iris recognition [J]. Machine Vision and Applications, 1996, 9 (1): 1 - 8.
- [3] DAVIES E R. Machine vision: studies from Anna university in the area of machine vision reported (machine vision inspection system for detection of leather surface defects) [J]. Journal of Engineering, 2019, [https://kns.cnki.net/kcms/detail/detail.aspx? FileName=SPQD5216761D12C51DCA1CB54BDD34A4245F&.DbName=GARJ2019](https://kns.cnki.net/kcms/detail/detail.aspx?FileName=SPQD5216761D12C51DCA1CB54BDD34A4245F&.DbName=GARJ2019).
- [4] HUANG C Y, HONG J H, HUANG E. Developing a machine vision inspection system for electronics failure analysis [J]. IEEE Transactions on Components, Packaging and Manufacturing Technology, 2019, 9 (9): 1 - 14.
- [5] DAVIES E R. Machine vision: theory, algorithms, practicalities [J]. CVGIP: Image Understanding, 1991, 54 (1): 1.
- [6] 郭 慧. 基于机器视觉的刨花板表面缺陷在线检测系统研究 [D]. 北京: 中国林业科学研究院, 2019.
- [7] 王昌书, 黄沿江, 张宪民, 等. 一种基于机器视觉的曲面玻璃划痕缺陷检测方法 [J]. 自动化技术与应用, 2020, 39 (1):

134-139.

- [8] 彭向前. 产品表面缺陷在线检测方法研究及系统实现 [D]. 武汉: 华中科技大学, 2008.
- [9] 江佳斌. 基于机器视觉的定位及缺陷识别智能检测技术与应用 [D]. 杭州: 浙江大学, 2020.
- [10] 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述 [J]. 数据采集与处理, 2016, 31 (1): 1-17.
- [11] 孙志军, 薛磊, 许阳明, 等. 深度学习研究综述 [J]. 计算机应用研究, 2012, 29 (8): 2806-2810.
- [12] 鲁越. 基于深度学习的手机玻璃缺陷分类检测 [D]. 郑州: 郑州大学, 2019.
- [13] 查云威, 陈志豪, 李伟朝. 基于改进 Faster Rcnnc 的手机屏幕缺陷检测方法研究 [J]. 计算机应用文摘, 2022, 38 (22): 3.
- [14] 金映谷, 张涛, 杨亚宁, 等. 基于深度学习的产品缺陷检测方法综述 [J]. 大连民族大学学报, 2020, 22 (5): 420-427.
- [15] 高如新, 任晓朵, 吴献, 等. 基于图像处理的手持屏幕缺陷检测系统研究 [J]. 测控技术, 2017, 36 (4): 26-30.
- [16] CHOI E, KIM J. Deep learning based defect inspection using the intersection over minimum between search and abnormal regions [J]. International Journal of Precision Engineering and Manufacturing, 2020, 21: 1-12.
- [17] CHEN Z, ZHA Y, WU Z, et al. Detection of mobile phone screen defect based on faster R-CNN fusion model [C] // 2021 China Automation Congress (CAC). IEEE, 2021: 6601-6606.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C] // Munich, Germany: Springer International Publishing, 2015: 234-241.
- [19] 孙文政. 基于深度学习和机器视觉的手机屏幕瑕疵检测方法研究 [D]. 济南: 山东大学, 2019.
- [20] 王春瑶, 陈俊周, 李炜. 超像素分割算法研究综述 [J]. (上接第 230 页)
- [19] DANELLJAN M, ROBINSON A, KHAN F S, et al. Beyond correlation filters; learning continuous convolution operators for visual tracking [J]. ArXiv, 2016, abs/1608.03773.
- [20] DANELLJAN M, BHAT G, KHAN F S, et al. ECO: Efficient convolution operators for tracking [C] // Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6931-6939.
- [21] VALMADRE J, BERTINETTO L, HENRIQUES J, et al. End-to-end representation learning for correlation filter based tracking [C] // Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5000-5008.
- [22] GAO L, LIU B, FU P, et al. Visual tracking via dynamic saliency discriminative correlation filter [J]. Applied Intelligence, 2021, 52: 5897-911.
- [23] WANG M, LIU Y, HUANG Z. Large margin object tracking 计算机应用研究, 2014, 31 (1): 6-12.
- [21] 宋熙煜, 周利莉, 李中国, 等. 图像分割中的超像素方法研究综述 [J]. 中国图象图形学报, 2015, 20 (5): 599-608.
- [22] ACHANTA R, SHAJI A, SMITH K, et al. SLIC Superpixels compared to state-of-the-art superpixel methods [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34 (11): 2274-2282.
- [23] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Las Vegas, USA: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016: 770-778.
- [24] 王文成, 蒋慧, 乔倩, 等. 基于 ResNet50 网络的十种鱼类图像分类识别研究 [J]. 农村经济与科技, 2019, 30 (19): 60-62.
- [25] 张珂, 冯晓晗, 郭玉荣, 等. 图像分类的深度卷积神经网络模型综述 [J]. 中国图象图形学报, 2021, 26 (10): 2305-2325.
- [26] 王元东. 基于 ResNet 模型的图像分类方法及应用研究 [D]. 南昌: 华东交通大学, 2019.
- [27] DUAN X. StegoPNet: image steganography with generalization ability based on pyramid pooling module [J]. IEEE Access, 2020, 8 (8): 195253-195262.
- [28] 王琳, 卫晨, 李伟山, 等. 结合金字塔池化模块的 YOLOv2 的井下行人检测 [J]. 计算机工程与应用, 2019, 55 (3): 133-139.
- [29] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C] // Las Vegas, USA: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881-2890.
- [30] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40 (4): 834-848.
- [24] MARVASTI-ZADEH S M, CHENG L, GHANEL-YAKHDAN H, et al. Deep learning for visual tracking: a comprehensive survey [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23 (5): 3943-68.
- [25] 严小乐. 基于相关滤波的视频运动目标跟踪算法研究 [D]. 南京: 南京航空航天大学, 2017.
- [26] STEPHEN B, NEAL P, ERIC C, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers [M]. Now, 2011.
- [27] WANG N, SHI J, YEUNG D Y, et al. Understanding and diagnosing visual tracking systems [C] // Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), 2015: 3101-3109.