

# 基于改进 PSPNet 的手机 LCD 屏幕表面缺陷检测

肖彬, 陈平华

(广东工业大学 计算机学院, 广州 510006)

**摘要:** 手机屏幕是智能手机的关键部件, 其品质优劣直接影响到用户的使用体验; 因此, 手机屏幕缺陷检测成为工业生产中的重要环节; 然而, 手机 LCD 屏幕的表面缺陷检测目前还存在检测精确度低、模型参数较多等问题, 无法满足实际工业生产需求; 为了解决这些问题, 对现有的缺陷检测算法和经典语义分割模型进行了研究, 提出一种基于改进 PSPNet 的手机 LCD 屏幕表面缺陷检测模型; 模型采用 MobileNetV3 作为特征提取网络, 有效减少了模型参数; 采用多尺度金字塔池化模块, 进一步整合多尺度上下文信息, 提高了模型的特征提取能力, 有效应对屏幕图像中缺陷尺寸微小、边界模糊、相同缺陷尺寸差异较大的问题; 同时, 通过引入注意力机制, 增强了模型的鲁棒性; 实验结果表明, 在 SQ、Mura、TP、Line 四种类型的手机 LCD 屏幕表面缺陷检测上, 改进后的模型准确度明显优于基线模型。

**关键词:** 手机 LCD 屏幕; 深度学习; 缺陷检测; PSPNet; 多尺度金字塔池化

## Surface Defect Detection of Mobile Phone LCD Screen Based on Improved PSPNet

XIAO Bin, CHEN Pinghua

(School of Computer Science, Guangdong University of Technology, Guangzhou 510006, China)

**Abstract:** The screen of mobile phone is a key component of smartphones, the quality of screen directly affects the experience of users. Therefore, the screen defect detection of mobile phone has become an important part of industrial production. However, the surface defect detection of mobile liquid crystal display (LCD) screens has the problems of low detection accuracy and a large number of model parameters, which can not meet actual industrial production needs. To solve this problems, this paper studies existing defect detection algorithms and classical semantic segmentation models, and proposes an improved mobile phone LCD screen defect detection model based on pyramid scene parsing network (PSPNet). The model adopts the MobileNetV3 as the feature extraction network, which effectively reduces the model parameters. the multi-scale pyramid pooling module is used to further integrate the multi-scale contextual information, improving the feature extraction ability of the model. It also effectively addresses the issues of small defect sizes, blurred boundaries, and significant differences in same defect size in screen images. Meanwhile, the attention mechanism is introduced to increase the robustness of the model. Experimental results show that the accuracy of the improved model is significantly better than that of other traditional semantic segmentation models in the surface defect detection of four LCD screens; SQ, Mura, TP, and Line.

**Keywords:** mobile phone LCD screen; deep learning; defect detection; PSPNet; multi-scale pyramid pooling

## 0 引言

随着互联网的日益普及, 信息的迅速扩散, 智能手机已经成为了人类最重要的通信手段。而手机屏幕是其中的核心部件, 它的质量将会影响到使用者的使用体验。目前, 液晶显示屏 (LCD, liquid crystal display) 是手机屏幕的主流选择, 与传统的屏幕材料相比, LCD 屏幕具备体积小、厚度适中、功耗低和安全性高等显著优势。然而由于 LCD 屏幕中需要注入的液晶分子数量较多, 并且对于电场控制系统的要求较高<sup>[1-2]</sup>, 所以生产 LCD 屏幕需要极其复杂、规范的工艺, 以当前的科技生产水平还无法避免缺陷的产

生。因此需要对制造出来的 LCD 屏幕进行严格、高效和快速的缺陷检测。

随着机器视觉的发展, 许多基于机器视觉的缺陷检测方式不断涌现<sup>[3-5]</sup>。文献 [6] 提出了基于图像处理的手​​机屏幕缺陷检测系统, 首先使用改进的开关中值滤波对图像去噪, 再通过傅里叶变换进行图像校正, 最后使用图像差分算法进行快速匹配, 实现缺陷特征的提取。文献 [7] 提出了基于邻域的背景差分算法, 在解决 LCD 屏幕亮度不均的同时能够有效突出 Mura 缺陷。并基于 MSER 提出了一种自适应阈值分割方法对屏幕进行缺陷分割, 准确率达到

收稿日期: 2024-03-22; 修回日期: 2024-04-28。

基金项目: 广东省重点领域研发计划项目 (2023B1111050010; 2020B0101100001)。

作者简介: 肖彬 (1993-), 女, 硕士研究生, 高级工程师。

陈平华 (1967-), 男, 教授, CCF 杰出会员。

引用格式: 肖彬, 陈平华. 基于改进 PSPNet 的手机 LCD 屏幕表面缺陷检测[J]. 计算机测量与控制, 2024, 32(9): 36-43.

95%以上。文献 [8] 提出了一种计算高效的基于回归诊断的 Mura 缺陷检测方法, 该方法利用线性模型估计输入图像的灰度数据, 通过计算 PRESS 残差滤除 Mura 区域, 再通过图像差分完成缺陷检测。文献 [9] 针对 TFT-LCD 屏幕提出了一套基于图像处理和计算机视觉的缺陷检测程序, 采用离散余弦变换和奇异值分解技术重建图像, 再通过图像差分完成缺陷检测。文献 [10] 针对 TFT-LCD 屏幕的 Mura 缺陷提出了一种基于多项式曲面拟合的缺陷检测技术, 首先将每个像素的灰度值假设为该像素二维坐标的函数, 再使用二元多项式对其进行曲面拟合。由于 Mura 缺陷面积较小, 拟合曲面可以看作是去除 Mura 缺陷后的背景模型。采用背景减除的方式即可将疑似 Mura 缺陷与背景区分开来。

尽管基于机器视觉的缺陷检测技术已显著超越了传统的人工检测手段, 但其在适用性、迁移性和鲁棒性方面依然存在明显不足, 越来越多的研究人员将目光投向了深度学习领域。文献 [11] 提出了一种针对手机屏幕的缺陷检测方法 Attention U-Net, 该方法先通过分类网络判断表面是否存在缺陷, 再对有缺陷的图像进行缺陷检测, 并通过加入注意力机制, 提升了检测的效率和精确度。文献 [12] 提出了一种适用于小样本学习的手机屏幕缺陷检测模型, 首先使用深层卷积生成对抗网络增强和生成缺陷样本, 然后对 FasterR-CNN 进行改进实现缺陷检测。文献 [13] 基于 Faster R-CNN 提出了一种手机屏幕缺陷检测模型, 首先使用边界平衡生成对抗网络对数据进行增强和生成, 然后将 FPN 和 ResNet 结合起来替换 Faster R-CNN 原始的特征提取网络, 并将 ROI 池化层替换成 ROI Align 层, 能够实现 99.43% 的准确率。文献 [14] 也提出了一种基于 Faster R-CNN 的手机屏幕缺陷检测模型, 使用 U-Net 和 ResNet 相结合作为模型的特征提取网络, 同时采用 K-means 聚类算法生成特定尺寸的初始锚点, 以提高分类的准确性和速度。可以看出, 针对手机屏幕和 LCD 屏幕的缺陷检测研究, 国内外研究人员均已取得了一定成果。但在手机 LCD 屏幕缺陷检测领域, 深度学习算法尚未得到广泛应用。因

而, 基于深度学习的手机 LCD 屏幕缺陷检测研究具有广泛前景和重大意义。

此外, 现有的基于深度学习的缺陷检测网络结构往往参数较多, 模型体积庞大, 不能很好地适应高效、轻便的工业应用需求。本文提出一种对手机 LCD 屏幕进行缺陷检测的轻量级卷积神经网络模型, 可应用于手机屏幕生产过程中的缺陷检测环节, 能够有效提升缺陷检测的准确率和效率, 进而提升工业生产的自动化水平。

## 1 PSPNet

金字塔场景解析网络 (PSPNet, pyramid scene parsing network)<sup>[15]</sup> 是一种典型的针对图像语义分割任务的网络结构, 该网络通过对多尺度的上下文信息地有效融合, 使语义分割的准确性和鲁棒性得到了显著提高。在 Pascal VOC 2012 数据集上, PSPNet 的 mIoU 准确率达到 85.4%。在更具挑战性的 Cityscapes 数据集上, PSPNet 也实现了 80.2% 的准确率。因此, 本文选择 PSPNet 作为手机屏幕缺陷检测的基础模型。

PSPNet 模型的网络结构主要由三部分组成: 特征提取网络、金字塔池化模块和输出端, 其网络结构图如图 1 所示。

1) 特征提取网络采用 ResNet50<sup>[16]</sup> 结构, 这是一种深度残差网络。残差结构的引入解决了深度网络中容易出现的性能退化的问题, 使得即使是拥有上千层深度的网络结构也能保持出色的表现。具体来说, 残差结构首先利用  $1 \times 1$  的卷积来进行维度的调整, 先升维后降维, 显著减少了要计算的参数量。其次使用跳跃连接, 将输入直接添加到输出中形成“残差”, 从而弥补卷积操作中可能损失的特征信息。这种结构让网络能够更容易地学习残差函数, 在增加深度的情况下不增加训练难度, 同时也可以更有效的进行特征信息的传递, 避免性能下降。残差结构如图 2 所示。

2) 金字塔池化模块 (PPM, pyramid pooling module)<sup>[17-18]</sup> 的目的是将不同尺度的特征进行融合, 以捕捉不同大小的上下文信息, 从而获取更全面的全局信息。PPM 包括 4 个不同尺度的池化层, 将输入图像分别划分为  $1 \times 1$ 、2

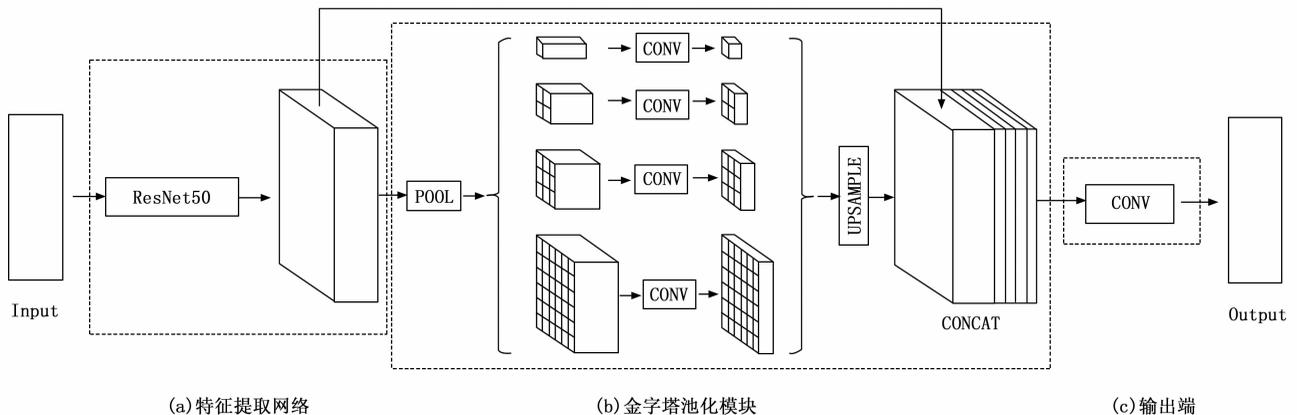


图 1 PSPNet 结构图

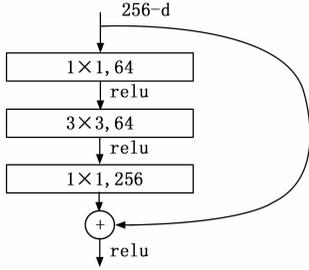


图 2 ResNet 的残差结构图

$\times 2$ 、 $3 \times 3$  和  $6 \times 6$  大小的子区域，对每个子区域进行平均池化，即可得到 4 张不同尺寸的特征图。每张特征图都利用  $1 \times 1$  的卷积降低通道数，再通过双线性插值的方式上采样到输入图像的尺寸，就能得到 4 张大小和输入图像一样、通道数为输入图像的  $1/4$  的特征图。最后，将这 4 张特征图与输入图像进行拼接，就能得到一张融合了多尺度上下文信息的复合特征图。

3) 在输出端，通过卷积操作对 PPM 生成的复合特征图进行解码，对每个像素点进行分类，获得最终的预测结果。

## 2 改进的 PSPNet

针对 PSPNet 对小目标检测的精确度不够高、模型参数较多等问题<sup>[19-22]</sup>，从 3 个方面对其进行改进，改进后的 PSPNet 网络结构如图 3 所示。

### 2.1 特征提取模块

PSPNet 所使用的骨干特征提取网络 (backbone) 为 ResNet50，该网络解决了深度神经网络中常见的梯度消失和性能退化的问题。残差模块的引入，使得网络在增加深度的同时也能保持良好的性能。尽管残差结构已经大大减

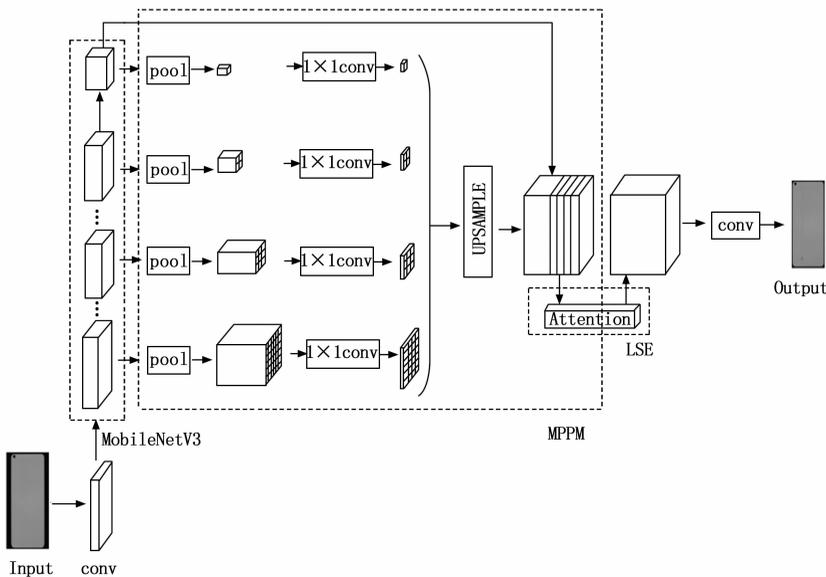


图 3 改进后的 PSPNet 示意图

少了计算参数量，但在工业应用中，还是追求更轻量化的模型。因此，本文选用 MobileNetV3<sup>[23]</sup> 作为 backbone 对 PSPNet 进行改进。

MobileNetV3 是图像分类领域的典型轻量级模型，该模型中的典型结构是倒残差结构。在倒残差结构中，使用  $1 \times 1$  卷积进行的是先升维再降维的操作，与残差结构刚好相反。这种设计虽然有助于增强网络的特征表示能力，但同时也会导致模型参数数量的增加。为了减轻这种影响并提高模型的效率，作者在倒残差结构中引入了深度可分离卷积来替代残差结构中的普通卷积。

深度可分离卷积<sup>[24]</sup> 是一种更为轻量级的卷积方式，它可以被看作是将标准的卷积操作分为两个步骤：首先是逐通道卷积，它独立地对每个输入通道采用不同的卷积核；然后是逐点卷积，它使用  $1 \times 1$  的卷积核对逐通道卷积的输出进行线性组合；如图 4 所示。这种分解方式显著减少了卷积操作中的参数量，提高模型的推理速度的同时保留足够的特征提取能力。

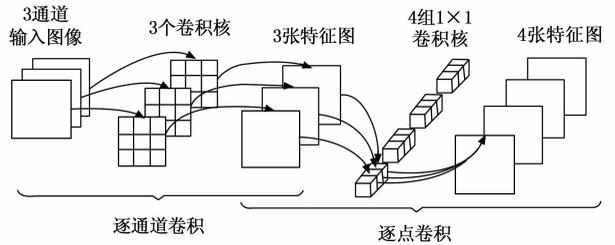


图 4 深度可分离卷积示意图

在残差结构中，所有结构块内部均进行了跳跃连接。而在 MobileNetV3 的倒残差结构中，只有结构块中深度可分离卷积的  $\text{stride}=1$  且输入图像和输出图像的通道数一致时，才进行跳跃连接。此外，部分结构块中引入了 SE 注意力机制，不同的结构块所用的激活函数也不一致，激活函数公式如表 1 所示，完整的结构块示意图如图 5 所示。

表 1 MobileNetV3 激活函数

缩写	激活函数	表达式
RE	ReLU6	$\min[\max(x, 0), 6]$
HS	h-sigmoid	$\frac{\text{ReLU6}(x+3)}{6}$

本文中 MobileNetV3 作为 backbone 模块，在原始模型的基础上做了一些调整，共包含 2 个卷积层和 15 个倒残差模块，完整的网络结构如表 2 所示。Exp size 栏指的是输入图像在倒残差结构中经过升维操作之后的通道数，#out 栏则指的是本单元经过降维操作后所输出图像的通道数。SE 栏指的是是否使用 SE 通道注意力机制，“ $\checkmark$ ”表示在对应 block 单元中加入了 SE 注意力机制，“-”则

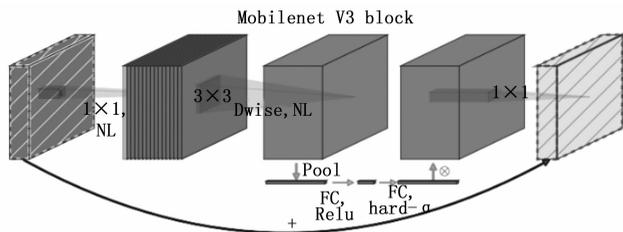


图 5 MobileNetV3 中完整结构块示意图

表示未使用注意力机制。NL 栏对应激活函数。在 backbone 中将提取中间 3 个倒残差模块输出的不同尺寸的特征图, 与最后的特征图一起作为 MPPM 模块的输入。

表 2 MobileNetV3 网络结构

Input	Operator	exp size	# out	SE	NL	S	d
512 <sup>2</sup> × 3	Conv, 3 × 3	—	16	—	HS	2	1
256 <sup>2</sup> × 16	Block, 3 × 3	16	16	—	RE	1	1
256 <sup>2</sup> × 16	Block, 3 × 3	64	24	—	RE	2	1
128 <sup>2</sup> × 24	Block, 3 × 3	72	24	—	RE	1	1
128 <sup>2</sup> × 24	Block, 5 × 5	72	40	✓	RE	2	1
64 <sup>2</sup> × 40	Block, 5 × 5	120	40	✓	RE	1	1
64 <sup>2</sup> × 40	Block, 5 × 5	120	40	✓	RE	1	1
64 <sup>2</sup> × 40	Block, 3 × 3	240	80	—	HS	2	1
32 <sup>2</sup> × 40	Block, 3 × 3	200	80	—	HS	1	1
32 <sup>2</sup> × 40	Block, 3 × 3	184	80	—	HS	1	1
32 <sup>2</sup> × 40	Block, 3 × 3	184	80	—	HS	1	1
32 <sup>2</sup> × 40	Block, 3 × 3	480	112	✓	HS	1	1
32 <sup>2</sup> × 112	Block, 3 × 3	672	112	✓	HS	1	1
32 <sup>2</sup> × 112	Block, 5 × 5	672	160	✓	HS	1	2
32 <sup>2</sup> × 160	Block, 5 × 5	960	160	✓	HS	1	2
32 <sup>2</sup> × 160	Block, 5 × 5	960	160	✓	HS	1	2
32 <sup>2</sup> × 160	Conv, 1 × 1	—	960	—	HS	1	1

### 2.2 多尺度金字塔池化模块 MPPM

为了清晰地捕捉手机 LCD 屏幕上的各种缺陷, 屏幕图像通常具备较高的分辨率, 而缺陷在屏幕图像中的占比较小。此外, 缺陷的形状和大小也各不相同, 部分缺陷的尺寸甚至存在显著差异, 且屏幕边缘位置也会存在不少缺陷, 这些都给缺陷检测的精确性带来了不小的挑战。针对上述问题, 提出了多尺度金字塔池化模块 (MPPM, multi-scale pyramid pooling module) 模块, 如图 6 所示, 针对性的聚合多尺度上下文信息, 提高缺陷检测的准确性。

在骨干特征提取网络中, 不同层次的网络负责提取不同类型的特征信息。浅层部分的网络经过的卷积层数量相对较少, 感受野偏小, 因此主要捕捉图像的局部特征信息, 包括边缘、纹理、颜色等。随着网络层数的加深, 每一次卷积运算都会进一步抽象整合之前的特征, 提炼出涵盖了图像整体结构和

语义内容的全局特征<sup>[25]</sup>。而在池化层中, 网络对全局特征和局部特征的捕捉能力是由池化窗口的大小决定的。较小的窗口可以捕捉到更细微的特征, 因为每个池化操作只涉及较小的图像区域, 使得网络能够更加细致地分析图像中的每一个细节, 这有助于保留更多的细节信息。相反, 较大的窗口在每个池化操作中能够覆盖较大的图像区域, 使得网络能够更好地捕捉到图像的宏观结构, 提取出更具有代表性的全局特征<sup>[26-28]</sup>。

因此, 在 MPPM 模块中, 对于从 backbone 输入的不同层次不同大小的特征图, 将分别使用不同大小的窗口对其进行池化操作。对于浅层网络获取的特征图, 为了更精细地捕捉局部细节和边缘信息, 采用较小的窗口进行池化操作; 深层网络获取的特征图则采用较大的窗口进行池化操作, 以便整合更多的空间信息, 捕捉全局特征。这种设计能够更针对性地捕获上下文信息, 扩大网络的感受野, 提高特征捕获能力, 从而提高缺陷检测的精确度。

MPPM 模块具体步骤如下:

- 1) 本文中特征金字塔的层数  $N=4$ , 故从 backbone 中提取 3 张层次不同的特征图和最后的 feature map 一起作为 MPPM 模块的输入。
- 2) 将  $N$  张特征图按层次从深到浅排列, 依次按照  $1 \times 1$ 、 $2 \times 2$ 、 $3 \times 3$ 、 $6 \times 6$  的大小划分并进行平均池化。
- 3) 将池化后的  $N$  张特征图通过  $1 \times 1$  卷积统一将通道数调整为 feature map 通道数的  $1/N$ 。
- 4) 利用双线性插值操作将步骤 3) 得到的特征图上采样到 feature map 的尺寸, 并将这  $N$  张特征图与 feature map 拼接在一起, 形成一个通道数为 feature map 两倍的融合了多尺度上下文信息的复合特征图。

### 2.3 通道注意力机制

注意力机制最初起源于对人类视觉的研究, 是一种模拟人类视觉和认知系统的方法。它赋予了神经网络感知适应的能力, 使其面对大量输入时, 能选择性地关注重要的信息, 提取关键特征。注意力机制通常作为一个关键模块嵌入神经网络之中, 它通过对输入特征图进行压缩和激励,

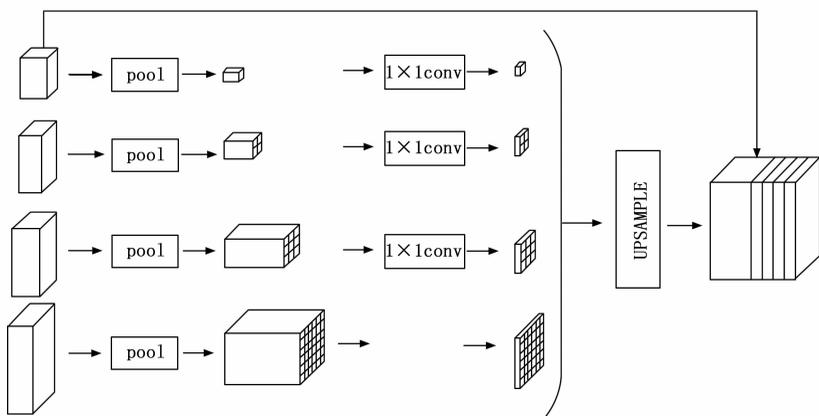


图 6 MPPM 模块示意图

动态调整网络的权重分配，确保网络能够捕捉到更为关键的信息，进而优化其性能。特别是在多尺度特征融合网络之后，融入注意力机制，能够增强目标信息的特征，抑制无效信息的影响，从而提升网络在复杂环境中的抗干扰能力。因此，将注意力机制模块加入到 MPPM 模块后面。考虑到工业应用所需的便利性，提出了一种简化的通道注意力机制 LSE (Light Squeeze-and-Excitation) 模块，如图 7 所示。

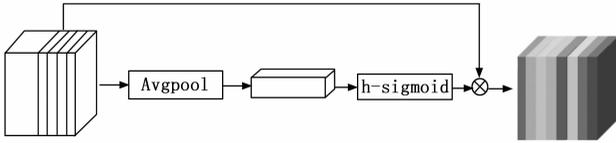


图 7 LSE 模块示意图

LSE 保留了通道注意力网络 (SENet, squeeze and excitation network)<sup>[29]</sup> 中压缩、激励和缩放的 3 个核心操作。在压缩阶段，沿着空间维度进行特征压缩，使用全局平均池化，将  $C \times H \times W$  大小的特征图压缩成一个  $1 \times 1 \times C$  的特征向量  $Z$ ，这个特征向量某种程度上拥有全局的感受野。该过程的数学表达式为：

$$z_c = F_{sq}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (1)$$

其中： $x_c$  表示  $X$  中的第  $c$  个特征图，后面公式中同理。

在激励阶段，为了利用压缩操作中汇聚的信息，使用 h-sigmoid 激活函数来为每个特征通道生成权重，其表达式如表 1 所示。最后在缩放阶段，将激励阶段获取的权重以逐通道相乘的方式赋到原始特征图上，完成对每个通道重要程度的标定。该过程的数学表达式为：

$$\tilde{x}_c = F_{scale}(x_c, s_c) = s_c \cdot x_c \quad (2)$$

### 3 实验结果与分析

#### 3.1 数据集

实验中的数据集来自广东省某手机屏幕生产企业，手机实体屏幕为 LCD 屏幕，本身不具有透明性。图像采集时需将屏幕置于工控机上，使用 PG 盒子给手机屏幕发送电信号以点亮屏幕，同时设置一个 LED 面阵光源打背光，两个 LED 条形光源进行低角度侧面打光。原始手机屏幕图像如图 8 所示。



图 8 原始手机屏幕图像

由于工业生产中使用的相机普遍分辨率较高，其所拍摄的原始手机屏幕图像的分辨率高达  $6\,576 \times 4\,384$  像素，这种高分辨率的图像如果直接输入到卷积神经网络中进行处理，将会耗费大量的内存和显存资源。相比之下，缺陷往往只有十几像素甚至更小，直接进行缺陷检测的效果往往也不尽如人意。因此，需要在实验前对原始图像进行切割，获得  $512 \times 512$  的图像。同时，为了提高模型的预测效果并防止出现过拟合的问题，切割后的图片还要进行数据增强和预处理操作，以提高数据集的数量和质量。实验中主要采取翻转、旋转、裁剪等数据增强操作，以及调节亮度饱和度、直方图均衡化和高斯模糊等预处理操作。

数据集中包含 SQ、Mura、TP、Line 四种类型的缺陷，SQ 缺陷指的是圆形、椭圆形或形状接近圆形的缺陷；Mura 缺陷指的是不规则的斑状缺陷；TP 缺陷指的是正方形缺陷，颜色有偏白色的，也有偏黑色的；Line 缺陷指的是线性缺陷，一般是贯穿整个屏幕的，有横向的也有纵向的，还有少量是断开的不连续的线缺陷。如图 9 所示。处理过后的数据集总共包含 20 091 张图片，将其按照 8 : 1 : 1 的比例拆分为训练集、验证集和测试集，具体分布如表 3 所示。

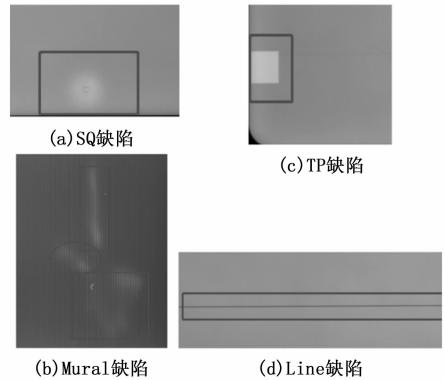


图 9 手机 LCD 屏幕缺陷示意图

表 3 数据集缺陷数量分布

数据集	SQ 缺陷	Mura 缺陷	TP 缺陷	Line 缺陷	无缺陷图片	总计
训练集	3 016	4 696	2 840	3 960	1 561	16 073
测试集	377	587	355	495	195	2 009
验证集	377	587	355	495	195	2 009

#### 3.2 实验配置与参数

实验所使用的计算机硬件设备为 Intel Core i5-7300HQ @ 2.50 GHz 处理器和一张 12 G 的 NVIDIA GeForce RTX3080 显卡。使用 Python 编程语言对 PSPNet 进行改进，并借助深度学习框架 Pytorch 来搭建实验模型。经过预处理后的输入图像尺寸为  $512 \times 512$ ，num\_classes 为 5；初始学习率设置为 .0.001，并采用余弦衰减的方式进行调整；为了避免模型在学习过程中陷入局部最小值，batch\_size 设置为 16，epoch 设置为 200，并使用 Adam 算法对模型进行优化。训练时利用骨干特征提取网络 MobileNetV3 的预

训练权重进行迁移学习, 避免权值太过随机, 影响特征提取效果。

实验分为两部分, 第一部分为改进的 PSPNet 模型与其他主流的语义分割模型的横向对比实验, 包括原始 PSPNet 模型、U-Net 模型和 DeepLabV3 模型。第二部分为纵向对比的消融实验, 用来验证本文中的所有改进对手机屏幕缺陷检测性能的影响。

### 3.3 评价指标

由于语义分割任务是对图像进行像素级的分类, 对像素预测的精确度进行计算的像素精度 (PA)、平均像素精度 (mPA) 是最基础的评价指标。而平均交并比 (mIoU) 计算的是预测结果与真实标签之间的相似度, 综合考虑了各个类别的预测准确性, 更能体现模型的综合性能。本文将采用 mPA 和 mIoU 作为主要评价指标。

mPA 是在 PA 基础上做的提升, 计算的是每一类像素正确分类的概率。其公式为:

$$mPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \quad (3)$$

mIoU 计算的则是每一类中真实值和预测值交集和并集的比例。对于具体的某一类别来说, 关于真实值和预测值的集合, 主要可以分为 4 种情况: TP (真正例) 代表预测为真且实际为真的情况, FP (假正例) 代表预测为真实际为假的情况, TN (真反例) 代表预测为假实际为假的情况, FN (假反例) 代表预测为假实际为真的情况。所以 mIoU 的公式为:

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN + FP + TP} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ij}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (4)$$

其中:  $i$  为真实值,  $j$  为预测值,  $p_{ij}$  表示  $i$  被预测为  $j$  的情况,  $k$  为类别数量。

### 3.4 对比实验

为了验证改进后的 PSPNet 模型在手机 LCD 屏幕缺陷检测领域的有效性, 在上文所述数据集上对其进行了训练和测试, 取得了较为良好的性能表现。该模型针对各类缺陷的 mPA 和 mIoU 指标如表 4 所示。

表 4 手机屏幕缺陷检测结果 %

Class	mPA	mIoU
SQ	82.5	79.6
Mura	64.6	61.3
TP	81.8	80.1
Line	87.2	84.8
Background	96.2	93.7

改进后的 PSPNet 与其他语义分割模型在手机 LCD 屏幕数据集上的各类缺陷检测性能对比如表 5 所示。与原始 PSPNet 模型相比, 该模型在 mPA 和 mIoU 性能指标上分别提升了 18.7% 和 18.5%。性能提升的原因主要在于输入

金字塔池化模块的特征图来自不同层次, 这些特征图中含有不同尺度的信息。经过不同尺度的池化操作后, 进一步增强了上下文语义信息, 有效提高了检测性能。此外, 针对拼接后的多通道特征图还引入了通道注意力机制加以强化, 使模型更加关注缺陷信息, 抗干扰能力得到加强。

表 5 各模型的精确度对比 %

Method	SQ	Mura	TP	Line	mPA	mIoU
PSPNet	67.2	48.6	65.8	64.7	63.8	61.4
U-Net	62.4	37.3	59.1	60.4	55.3	53.8
DeepLabV3	68.4	52.4	58.3	65.4	64.5	61.7
Ours	78.6	62.1	78.6	86.4	82.5	79.9

与 U-Net 模型相比, 该模型在 mPA 和 mIoU 性能指标上分别提升了 27.2 和 26.1 个百分点。与 DeepLabV3 模型相比, 该模型在 mPA 和 mIoU 性能指标上分别提升了 18 和 18.2 个百分点。可以看出, 虽然 U-net 模型和 DeepLabV3 模型结构上都注重上下文信息的融合, 但在手机 LCD 屏幕的缺陷检测方面存在性能下降的现象, 这说明其算法结构在检测 LCD 屏幕的各类缺陷上适用性不强。

### 3.5 消融实验

为了验证本文所提出的 3 个改进模块的有效性, 在上文所述数据集上进行了 6 种方案的消融实验, 各方案的性能对比如表 6 所示。方案一直接以原始 PSPNet 模型为框架; 方案二将 PSPNet 中的骨干特征提取网络替换为轻量级网络 MobileNetV3; 方案三在方案二的基础上引入了多尺度金字塔池化模块。后 3 种方案则主要对注意力机制进行对比实验, 方案四加入了简化后的 SE 注意力模块 LSE, 是 PSPNet 的完整改进版; 方案五采用原始 SE 注意力机制; 方案六采用卷积块注意力模块 (CBAM, convolutional block attention module)<sup>[30]</sup>。

表 6 各模块消融实验性能对比 %

Method	mPA	mIoU
ResNet50+PPM	63.8	61.4
MobileNetV3+PPM	67.2	65.8
MobileNetV3+MPPM	76.4	75.7
MobileNetV3+MPPM+LSE	82.5	79.9
MobileNetV3+MPPM+SE	82.7	80.4
MobileNetV3+MPPM+CBAM	82.8	80.3

根据消融实验结果可以看出, 方案二中使用 MobileNetV3 结构替换 ResNet50 结构后, 较方案一在 mPA 和 mIoU 性能指标上分别提升了 3.4% 和 4.4%。尽管性能提升并不显著, 但 MobileNetV3 作为轻量级网络模型, 与 ResNet50 相比, 其参数量明显减少, 大幅降低了计算复杂度, 显著提升了推理速度。同时, 其相对简单的网络结构, 能够在训练和优化的过程中更快的找到最佳的参数配置, 提高训练效率。

方案三基于方案二引入了多尺度金字塔池化模块

(MPPM), 相比方案二在  $mPA$  和  $mIoU$  性能指标上分别提升了 9.2% 和 9.9%。通过将不同层次的特征图输入金字塔池化模块, 进行多尺度的池化操作, MPPM 模块综合保留了浅层特征的细节信息和深层特征的语义信息, 让网络能够充分捕捉到多尺度的上下文信息。这一改进显著提高了手机屏幕图像缺陷检测的准确性, 尤其是针对缺陷像素较小、大小和形状多样性问题有了更好的解决方案。

方案四为本文所提出的完整模型框架, 基于原始 PSP-Net 替换了骨干特征网络, 加入了 MPPM 模块和 LSE 模块, 较方案三在  $mPA$  和  $mIoU$  性能指标上分别提升了 6.1% 和 4.2%。在 MPPM 模块中, 进行了多尺度特征融合, 得到一张多通道的复合特征图。该特征图不仅包含了输入图像中不同尺度的信息, 还融合了来自不同层级的特征信息, 具有丰富的特征表示能力。在此处引入注意力机制, 为每个通道赋予不同的权重, 可以使模型聚焦于关键特征信息, 忽略无关信息, 进一步提升缺陷检测的准确度。

方案五采用 SE 注意力机制, 较方案四在  $mPA$  和  $mIoU$  性能指标上分别提升了 0.2% 和 0.5%; 方案六引入 CBAM 模块, 较方案四在  $mPA$  和  $mIoU$  性能指标上分别提升了 0.3% 和 0.4%。尽管方案五与方案六在性能方面有所提升, 但提升幅度相对较小, 且伴有更多的计算量和参数量。因此, 在工业现场检测的实际需求背景下, 方案四更具适用性。

#### 4 结束语

伴随手机普及率的持续提升, 手机 LCD 屏幕的生产规模逐步扩大, 使得手机屏幕缺陷检测环节在工业制造业中愈发关键。然而, 由于缺陷尺寸较小且形状各异等因素, 当前的缺陷检测精度尚待提高。因此, 本文提出了一种基于 PSPNet 的改进框架来解决这一难题。在改进框架中, 使用 MobileNetV3 作为骨干特征提取网络, 这显著降低了模型的计算负担和参数规模, 进而减少了对设备配置的要求, 使得在工业环境中部署应用更加方便。同时, 文中提出的 MPPM 模块, 使得模型能够更好的整合多尺度上下文信息, 提高了模型的特征提取能力, 增强了对小目标检测的精确度。此外, 针对 MPPM 模块输出的多通道复合特征图, 提出了 LSE 模块对其进行注意力加权, 以提高关键特征信息的权重, 降低无效信息的干扰, 进一步提高模型检测的准确率。在手机 LCD 屏幕数据集上的实验结果表明, 本文提出的改进措施可以有效提高手机屏幕缺陷检测的效果。

当然, 该框架还存在优化改进的空间。目前的检测效果虽有提升, 但是 Mura 缺陷的检测精度还相对较低, 在后续的工作中可以考虑采取一些预处理方式突出图像特征信息, 以提升对 Mura 缺陷的检测效果。此外, 还可以在不显著增加参数量的情况进一步优化注意力机制, 从而进一步提升模型的缺陷检测效果。

#### 参考文献:

[1] CHERIE R, KAGAN, PAUL ANDRY. Thin-film transistors

[M]. CRC Press, 2003, 16 - 26.

- [2] 廖 苗, 刘毅志, 欧阳军林, 等. 基于自适应局部增强的手机 TFT-LCD 屏 Mura 缺陷自动检测 [J]. 液晶与显示, 2018, 33 (6): 475 - 482.
- [3] 秦 钢. 基于图像处理的液晶屏幕缺陷检测 [D]. 广州: 广东工业大学, 2017.
- [4] 张清蓉, 陈龙灿, 南亮生. 基于机器视觉的 LCD 屏点缺陷检测技术研究 [J]. 南方农机, 2022, 53 (2): 143 - 145.
- [5] 熊召龙. 基于机器视觉的辐射液晶屏缺陷检测及修复系统的研究与开发 [D]. 广州: 华南理工大学, 2017.
- [6] 高如新, 任晓东, 吴 猷, 等. 基于图像处理的手机屏幕缺陷检测系统研究 [J]. 测控技术, 2017, 36 (4): 29 - 31.
- [7] 钱基德, 陈 斌, 钱基业, 等. 基于机器视觉的液晶屏 Mura 缺陷检测方法 [J]. 计算机科学, 2018, 23 (5): 52 - 55.
- [8] FAN S K S, CHUANG Y C. Automatic detection of Mura defect in TFT-LCD based on regression diagnostics-ScienceDirect [J]. Pattern Recognition Letters, 2010, 31 ( 15): 2397 - 2404.
- [9] YANG Y B, LI N, ZHANG Y. Automatic TFT-LCD mura detection based on image reconstruction and processing [C] // IEEE Third International Conference on Consumer Electronics-Berlin, 2013.
- [10] 张 昱. 基于机器视觉的 TFT-LCD 屏 mura 缺陷检测技术研究 [D]. 哈尔滨: 哈尔滨工业大学, 2006.
- [11] 李智勇. 基于深度学习的手机屏幕缺陷检测技术研究 [D]. 哈尔滨: 哈尔滨工业大学, 2020.
- [12] LV Y, MA L, JIANG H. A mobile phone screen cover glass defect detection MODEL based on small samples learning [C] // 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), IEEE, 2019.
- [13] WANG T, ZHANG C, DING R, et al. Mobile phone surface defect detection based on improved faster R-CNN [C] // International Conference on Pattern Recognition, IEEE, 2021. DOI:10.1109/ICPR48806.2021.9412119.
- [14] CHEN Z, ZHA Y, WU Z, ZENG D. Detection of mobile phone screen defect based on faster R-CNN fusion model [C] // China Automation Congress (CAC), 2021.
- [15] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [J]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, 6230 - 6239.
- [16] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2016.
- [17] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37 (9): 1904 - 1916.
- [18] DUAN X, WANG W, LIU N, et al. StegoPNet: image steganography with generalization ability based on pyramid pooling module [J]. IEEE Access, 2020, 8: 195253 - 195262.
- [19] REN S, LIU X, LIU H, et al. Cultivated land segmentation

- of remote sensing image based on pspnet of attention mechanism [J]. *International Journal of Robotics & Automation*, 2022 (1): 37 - 46.
- [20] DENG Z, ZHANG K, SU B, et al. Classification of breast cancer based on improved PSPNet [J]. *2021 IEEE/ACIS 6th International Conference on Big Data, Cloud Computing, and Data Science (BCD)*, 2021: 86 - 90.
- [21] YANG C, GUO H, YANG Z. A method of image semantic segmentation based on PSPNet [J]. *Mathematical Problems in Engineering*, 2022: 77 - 82.
- [22] ZHU X L G. Coronary angiography image segmentation based on PSPNet [J]. *Computer Methods and Programs in Biomedicine: An International Journal Devoted to the Development, Implementation and Exchange of Computing Methodology and Software Systems in Biomedical Research and Medical Practice*, 2021, 200 (1): 548 - 550.
- [23] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C] // *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, 2020.
- [24] CHOLLET F. Xception: deep learning with depthwise separable convolutions [J]. *Computer Vision and Pattern Recognition*, 2016, ArXiv: 1610.02357.
- [25] IAN GOODFELLOW, YOSHUA BENGIO, AARON COURVILLE. *Deep Learning* [M]. CRC Press, 2018, 21 - 26.
- [26] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2012, 25 (2): 672 - 679.
- [27] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [J]. *IEEE Computer Society*, 2014.
- [28] LIN M, CHEN Q, YAN S. Network in network [C] // *ICLR*, 2014.
- [29] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 42 (8): 201 - 220.
- [30] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module [C] // *Proceedings of the European Conference on Computer Vision*, 2018.
- (上接第 14 页)
- [3] 胡均平, 黄 强, 张洪伟, 等. 基于改进 YOLOv3 模型的车辆前方路面坑洼检测 [J]. *传感器与微系统*, 2022 (12): 130 - 133.
- [4] 焦双健, 杜福君. 基于深度学习的路面坑洼检测系统设计 [J]. *单片机与嵌入式系统应用*, 2022 (7): 10 - 13.
- [5] 戚玲珑, 高建瓴. 基于改进 YOLOv7 的小目标检测 [J]. *计算机工程*, 2023, 49 (1): 41 - 48.
- [6] 毛 剑. 基于 YOLOV7 蒸馏模型的道路缺陷检测算法研究 [J]. *现代制造技术与装备*, 2023, 59 (2): 222 - 224.
- [7] 叶嘉欣. 基于深度学习的两阶段目标检测算法综述 [J]. *互联网周刊*, 2023 (5): 16 - 18.
- [8] 韩岩江, 王 伟, 王峰萍. 融合坐标注意力和 BiFPN 的 YOLOv5s 交通标志检测方法 [J]. *国外电子测量技术*, 2022, 41 (11): 170 - 179.
- [9] SUN Y X, ZHANG Y J, WEI Z H, et al. A classification and location of surface defects method in hot rolled steel strips based on YOLOV7 [J]. *Metalurgija*, 2023, 62 (2): 240 - 242.
- [10] TANG F, YANG F, TIAN X Q. Long-distance person detection based on YOLOv7 [J]. *Electronics*, 2023, 12 (6): 1502.
- [11] LIU K Y, SUN Q, SUN D M, et al. Underwater Target Detection Based on Improved YOLOv7 [J]. *Journal of Marine Science and Engineering*, 2023, 11 (3): 677.
- [12] WU H X, ZHAO X, HAN G H, et al. MWSR-YLCA: Improved YOLOv7 embedded with attention mechanism for nasopharyngeal carcinoma detection from MR images [J]. *Electronics*, 2023, 12 (6): 1352.
- [13] ZHAO L L, ZHU M L. MS-YOLOv7: YOLOv7 based on multi-scale for object detection on UAV aerial photography [J]. *Drones*, 2023, 7 (3): 188.
- [14] YUAN WENAN. Accuracy comparison of YOLOv7 and YOLOv4 regarding image annotation quality for apple flower bud classification [J]. *AgriEngineering*, 2023, 5 (1): 413 - 424.
- [15] LAI Y H, MA R J, CHEN Y, et al. A pineapple target detection method in a field environment based on improved YOLOv7 [J]. *Applied Sciences*, 2023, 13 (4): 2691.
- [16] LEE SANGYUM, LE TRI HO MINH, KIM YEONG-MIN. Prediction and detection of potholes in urban roads: Machine learning and deep learning based image segmentation approaches [J]. *Developments in the Built Environment*, 2023, 13: 100 - 109.
- [17] SATTI SATISH KUMAR, K. SUGANYA DEVI, MADDULA PRASAD, et al. Unified approach for detecting traffic signs and potholes on Indian roads [J]. *Journal of King Saud University-Computer and Information Sciences*, 2022, 34 (10PB).
- [18] SATTI SATISH KUMAR, DEVI K. Suganya, Dhar Prasenjit, Srinivasan P. Detecting potholes on Indian roads using Haar feature-based cascade classifier, convolutional neural network, and instance segmentation [J]. *Soft Computing*, 2022, 26 (18): 9141 - 9153.
- [19] JIANG P, ERGU D, LIU F, et al. A review of yolo algorithm developments [J]. *Procedia Computer Science*, 2022, 199: 1066 - 1073.
- [20] BHAVANA K, MUNAPPA S, BHAVANI K D, et al. Automatic pothole and humps on roads detection and notification alert [C] // *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*. IEEE, 2023: 1 - 6.
- [21] INDRIYANI T, UTOYO M I, RULANINGTYAS R. Comparison of image smoothing methods on potholes road images [J]. *Journal of Physics: Conference Series*, 2020, 1477 (5): 052 - 056.