

太阳能电池片图像校正与表面缺陷检测^{*}

朱栋 胡伟笑 赵腾

(常州大学微电子与控制工程学院 常州 213159)

摘要: 针对太阳能电池片图像的透视失真与表面缺陷检测问题,提出一种基于虚拟相机的太阳能电池片图像的透视校正方法和改进 YOLOv5s 的神经网络模型。首先,根据相机外参构建水平姿态的虚拟相机,建立原图与虚拟相机的透视映射关系,以实现原图的透视校正。然后,采用动态头部来提高 YOLOv5s 头部的表示能力,并在 C3 模块的瓶颈处加入感受野增强模块 RFI 来提高小目标感受野。最后,将 YOLOv5s 的定位 loss 与 NWD loss 进行融合来弥补小目标位置偏差。实验结果表明,基于虚拟相机的透视校正,其效果明显优于传统方法且运行时间更短;同时改进后的 YOLOv5s 模型对比 YOLOv5s、YOLOv7、YOLOv8 平均精度分别提高 6.1%、27.7%、1.1%,对太阳能电池片表面质量检测具有实际应用价值。

关键词: 太阳能电池片;透视校正;虚拟相机;YOLOv5s;缺陷检测

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** G642

Image calibration and surface defect detection of solar photovoltaic cells

Zhu Dong Hu Weixiao Zhao Teng

(College of Microelectronics and Control Engineering, Changzhou University, Changzhou 213159, China)

Abstract: Aiming at image perspective distortion and surface defect detection of solar photovoltaic cells, a method based on a virtual camera for perspective correction and an improved YOLOv5s neural network model for defect detection are proposed. Firstly, a virtual camera with a horizontal orientation is constructed based on camera extrinsics to establish a perspective mapping relationship between the original image and the virtual camera, by which perspective correction of the original image is achieved. Secondly, a dynamic head is employed to enhance the representation capacity of the YOLOv5s head, and a receptive field expansion (RFI) module is added into the bottleneck of the C3 module to enhance the receptive field for small targets. Finally, the localization loss of YOLOv5s is fused with the normalized weighted distance (NWD) loss to compensate for the positional deviation of small targets. Experimental results demonstrate that the perspective correction based on the virtual camera can achieve significant improvements in correction effectiveness with shorter runtime. Moreover, the average accuracy of the improved YOLOv5s model can be increased up to 6.1%, 27.7%, and 1.1% than YOLOv5s, YOLOv7, and YOLOv8 respectively, which exhibits the practical value in surface quality inspection of solar photovoltaic cells.

Keywords: solar photovoltaic cell; perspective correction; virtual camera; YOLOv5s; defect detection

0 引言

太阳能作为重要的新型能源之一,为日常生活和工业生产提供可再生能源,可缓解能源紧缺问题^[1-2]。在太阳能电池片生产过程中,常见的表面缺陷,如划痕、黑点、无电和叉隐等 4 种缺陷,会影响电池片的转换效率和使用寿命。目前太阳能电池片的主流检测方法仍然是基于机器视觉和计算机视觉的检测方法^[3-4],如鲁东林等^[5]提出一种改进

Faster-RCNN (regions with convolutional neural network features) 的太阳能电池片缺陷检测方法,以 Faster R-CNN 为框架融入了基于 FPN (feature pyramid network, FPN) 的多尺度检测网络实现对太阳能电池片缺陷的检测,取得了较好的检测效果。王宪保等^[6]提出了深度置信网络 (deep belief network, DBN),通过训练该网络将它应用于太阳能电池片表面缺陷检测中。张会林等^[7]提出一种基于改进 YOLOv4 (you only look once version 4, YOLOv4) 的太

能电池片表面缺陷检测方法, 引入密集主干算法改进原模型的主干网络, 改善了算法的表面缺陷检测效果。王淑青等^[8]提出一种轻量化 YOLOX-S 检测模型, 以 YOLOX-S (you only look once x, YOLOX) 模型为基础采用轻量化网络 MobileNetV3 优化主干, 改进后的检测模型平均精度均值可达 97.6%。

为了快速检测太阳能电池片表面缺陷, 需要使用多个副相机来捕捉电池片的图像来对主相机图像进行补充。然而, 由于副相机的光轴与电池片的放置平台非垂直关系, 副相机捕捉到的电池片图像可能存在一定倾斜角度的形变, 影响检测精确性。透视校正方法主要有控制点变换法^[9]、双灭点法^[10]等。控制点变换法利用原图及目标图中 4 个及以上特征点对建立线性关系计算出单应性矩阵, 后对失真原图透视变换。Kim 等^[11]利用该方法在多种不同场景对图像透视校正。通常使用 Hough 变换、特征匹配等方法检测原图特征点, 且利用实际物理距离及像素精度估计目标图特征点。该方法对于特征明显或特定场景的图像校正效果较好, 对于背景更复杂的图像校正效果欠佳。双灭点法首先进行灭点估计, 然后通过双灭点模型分别校正水平及垂直灭点实现透视校正。Clark 等^[12]使用此方法对文本、车辆等进行校正, 此方法运用透视投影模型进行校正, 准确度较好。但通常也需要首先利用直线检测、特征匹配等方法估计灭点, 使用场景也受到限制, 且计算复杂不适用于对校正速度有要求的场景。Deng 等^[13]提出了一种并行卷积神经网络架构消除图像中的透视失真, 某些场景校正效果较好, 但前期需要大量数据图像进行训练, 且文中说明了该方法无法正确校正具有重复纹理或有杂乱物体图像, 故此方法对于具有密集栅线的太阳能电池片图像不适用。

针对以上问题, 本文第 1 章提出一种基于张正友虚拟相机的太阳能电池片图像校正方法, 该方法通过相机外参求解相机坐标系变换矩阵, 构建水平姿态的虚拟相机, 建立原图到虚拟相机下的透视映射关系, 避免控制点变换法造成图像质量的损失和变形效果的不自然, 从而获得没有透视的图像以提高缺陷检测精度。与传统控制点变换法进行对比, 实验结果表面, 本文提出的方法校正结果更加准确, 校正时间明显缩短且更具鲁棒性。第 2 章提出一种改进的 YOLOv5s (you only look once version 5, YOLOv5) 太阳能电池片表面缺陷检测方法, 通过添加动态头部 dyhead (dynamic head, Dyhead) 和感受野增强模块 RFI (receptive field improvement, RFI) 以及采用 NWD (normalized wasserstein distance, NWD) Loss 和 SE (squeeze and excitation, SE) 注意力模块等方法, 用以提高目标检测头部的表示能力和缺陷检测的准确性。第 3 章进行多组实验并对实验结果对比分析。

1 基于虚拟相机的图像透视校正

1.1 轴旋转角度的求解

通过标定得到的相机外参, 包括世界坐标系到相机坐

标系变换的旋转矩阵 \mathbf{R} 和平移向量 \mathbf{T} , 分别如式(1)和(2)所示。

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (1)$$

$$\mathbf{T} = [t_1 \quad t_2 \quad t_3] \quad (2)$$

式中: $r_{11}, r_{12}, r_{13}, r_{21}, r_{22}, r_{23}, r_{31}, r_{32}, r_{33}$ 为相机外参旋转矩阵 \mathbf{R} 中的参数; t_1, t_2, t_3 为相机外参平移向量 \mathbf{T} 中 X、Y、Z 轴的平移分量。其坐标旋转是一个正向旋转的过程, 先分别绕 Z、Y、X 轴旋转再平移得到相机坐标系。X、Y、Z 3 个坐标轴旋转 α, β, θ 角度的旋转分量为 $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z$, 所以旋转矩阵 \mathbf{R} 为:

$$\mathbf{R} = \mathbf{R}_x \mathbf{R}_y \mathbf{R}_z \quad (3)$$

即可得式(4)。

$$\mathbf{R} = \begin{bmatrix} \cos\beta\cos\theta & \cos\beta\sin\theta & -\sin\beta \\ -\cos\alpha\sin\theta + \sin\alpha\sin\beta\cos\theta & \cos\alpha\cos\theta + \sin\alpha\sin\beta\sin\theta & \sin\alpha\cos\beta \\ \sin\alpha\sin\theta + \cos\alpha\sin\beta\cos\theta & -\sin\alpha\cos\theta + \cos\alpha\sin\beta\sin\theta & \cos\alpha\cos\beta \end{bmatrix} \quad (4)$$

根据式(1)和(4), 可以分别解出 α, β 为:

$$\begin{cases} \alpha = \tan^{-1}(r_{23}/r_{33}) \\ \beta = -\sin^{-1}r_{13} \end{cases} \quad (5)$$

1.2 虚拟相机的构建

若实现透视校正需将原图转换为正俯视图, 如图 1 所示, $O_c - X_c Y_c Z_c$ 为原相机坐标系, 将其旋转为 $O'_c - X'_c Y'_c Z'_c$ 。使相机光轴(Z 轴)垂直于物体平面, 此时获得的图像即为俯视图, 同时要保证光轴与物体平面的交点不变, 则需要将相机坐标系再进行平移为 $O''_c - X''_c Y''_c Z''_c$ 。此过程即从原相机 A 旋转到相机 B, 最终平移至虚拟目标相机 C。

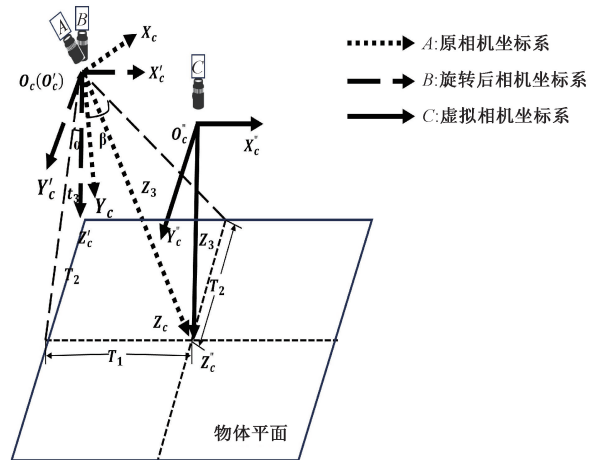


图 1 相机坐标系变换示意图

因世界坐标系的 Z 轴在标定时便设定与物体平面垂直, 所以只需将原相机坐标系先沿 X 轴反方向旋转 α 度, 再沿 Y 轴反方向旋转 β 度, 便可使相光轴与世界坐标系的 Z 轴平行, 从而垂直于物体平面。又因设定时世界坐标系为左手坐标系而相机坐标系为右手坐标系, 所以旋转后的

相机坐标系 Y 和 Z 轴取反方向,从而得到相机坐标系旋转变换关系如式(6)所示。

$$\begin{bmatrix} X'_c \\ Y'_c \\ Z'_c \end{bmatrix} = \begin{bmatrix} \cos\beta & \sin\alpha \sin\beta & \cos\alpha \sin\beta \\ 0 & -\cos\alpha & \sin\alpha \\ \sin\beta & -\sin\alpha \cos\beta & -\cos\alpha \cos\beta \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (6)$$

式中: Z 为坐标系某点在旋转后坐标系的坐标, (X_c, Y_c, Z_c) 为原相机坐标系某点。

图 1 中, t_3 为世界坐标系变换到相机坐标系变换的 Z 轴平移量,即原相机原点到物体平面的垂直距离, Z_3 为原相机坐标系中光心与物体平面沿光轴的距离。设 T_1, T_2, T_3 为 X'_c, Y'_c, Z'_c 方向上的平移距离,由图中坐标关系可建立方程解出相机变换的平移量如式(7)所示。

$$\begin{cases} T_1 = t_3 \tan\beta \\ T_2 = t_3 \tan\alpha \\ T_3 = Z_3 - t_3 = \sqrt{T_1^2 + T_2^2 + T_3^2} - t_3 \end{cases} \quad (7)$$

综上可建立原相机坐标系到目标相机坐标系的变换关系,以实现相机视角的变换。将变换关系写成齐次形式,如式(8)所示。

$$\begin{bmatrix} X''_c \\ Y''_c \\ Z''_c \end{bmatrix} = \begin{bmatrix} \cos\beta & \sin\alpha \sin\beta & \cos\alpha \sin\beta & T_1 \\ 0 & -\cos\alpha & \sin\alpha & T_2 \\ \sin\beta & -\sin\alpha \cos\beta & -\cos\alpha \cos\beta & T_3 \\ & & & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (8)$$

式中: (X''_c, Y''_c, Z''_c) 为原相机坐标系某点在虚拟相机坐标系的坐标。

1.3 图像的透视校正

由标定相关知识可知,三维世界坐标系转换到相机坐标系的关系如式(9)所示。

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (9)$$

式中: (X_c, Y_c, Z_c) 为该点的原相机坐标, (X_w, Y_w, Z_w) 为世界坐标系某点。三维相机坐标系转换到二维图像坐标系的关系如式(10)所示。

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (10)$$

式中: (u, v) 为原相机坐标, f_x, f_y 分别为 X 和 Y 方向上的像素焦距, u_0, v_0 分别为图像主点在 X 和 Y 方向上的偏移距离, (X_c, Y_c, Z_c) 对应的无畸变失真的原图像坐标。于是,原图像点转换至目标图点的透视映射矩阵 P 如式(11)所示。

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \quad (11)$$

本文透视校正方法在相机标定并畸变校正之后进行,所以基于张正友标定法所取的棋盘格图像计算透视映射矩阵,透视映射矩阵得到后需要利用它对太阳能电池片图像进行透视变换,其变换关系如式(12)所示。

$$\begin{cases} u' = \frac{p_{11}u + p_{12}v + p_{13}}{p_{31}u + p_{32}v + p_{33}} \\ v' = \frac{p_{21}u + p_{22}v + p_{23}}{p_{31}u + p_{32}v + p_{33}} \end{cases} \quad (12)$$

式中: (u', v') 为透视校正后图像点二维坐标。

1.4 图像校正结果分析

某张太阳能电池片原图、传统方法校正后图像及本文方法校正后图像分别如图 2(a)~(c)所示,对比可以看出本文提出的校正方法效果更好。

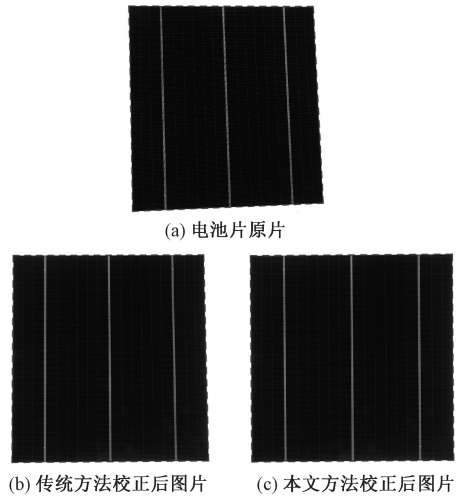


图 2 原图及校正后的图片

本文在 YOLOv5s 不添加任何功能模块的情况下,完成了太阳能电池片表面缺陷数据集校正前和校正后的对比实验,结果如表 1 所示。

表 1 电池片数据集校正前后实验对比

类别	准确率/%	召回率/%	mAP/%
校正前	73.6	75.0	80.4
校正后	76.9	81.5	83.9

由表 1 可知本文的图像校正效果良好,准确率提高 3.3%、召回率提高 6.5%、mAP 提高了 3.5%。

因为太阳能电池片是标准的正四边形,所以检测校正后电池片轮廓 4 条边并计算出临边平均夹角与 90° 的差值,可通过差值对比本文与传统方法的校正效果,结果如图 3 所示。

由图 3 夹角测试结果可以明显看出,本文透视校正效果更好,经计算传统方法校正后电池片四边轮廓平均夹角误差在 1.669 2°,标准差为 0.218 7°;而本文的方法透视校正后平均夹角误差为 0.497 3°,标准差为 0.057 5°,可看出本文方法夹角误差波动较小,更具鲁棒性。所以本文提出

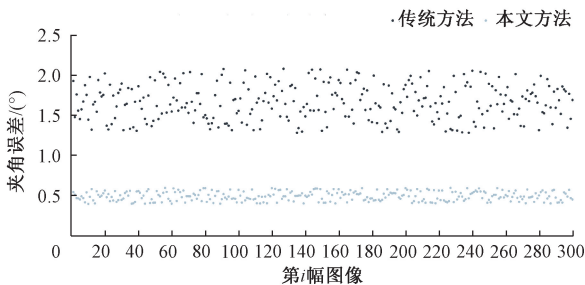


图 3 校正效果测试对比

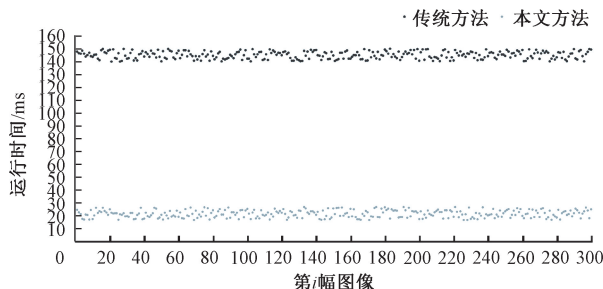


图 4 校正运行时间对比

的方法可很好的满足光伏组件自动化系统中对太阳能电池片检测的要求。

此外,对两种透视校正方法的运行时间也进行了对比测试,结果如图 4 所示。

由图 4 可知传统方法运行时间在 145 ms 左右,而本文透视校正方法仅为 25 ms 左右,比传统方法快了 120 ms 左右,可极大缩短对电池片的检测时间。

2 检测模型原理与优化

2.1 检测模型原理

太阳能电池片表面缺陷检测方法采用 YOLOv5s 卷积神经网络^[14],该结构简单易于学习、并且具有高实时性等特点。本文的太阳能电池片表面缺陷检测结构如图 5 所示。

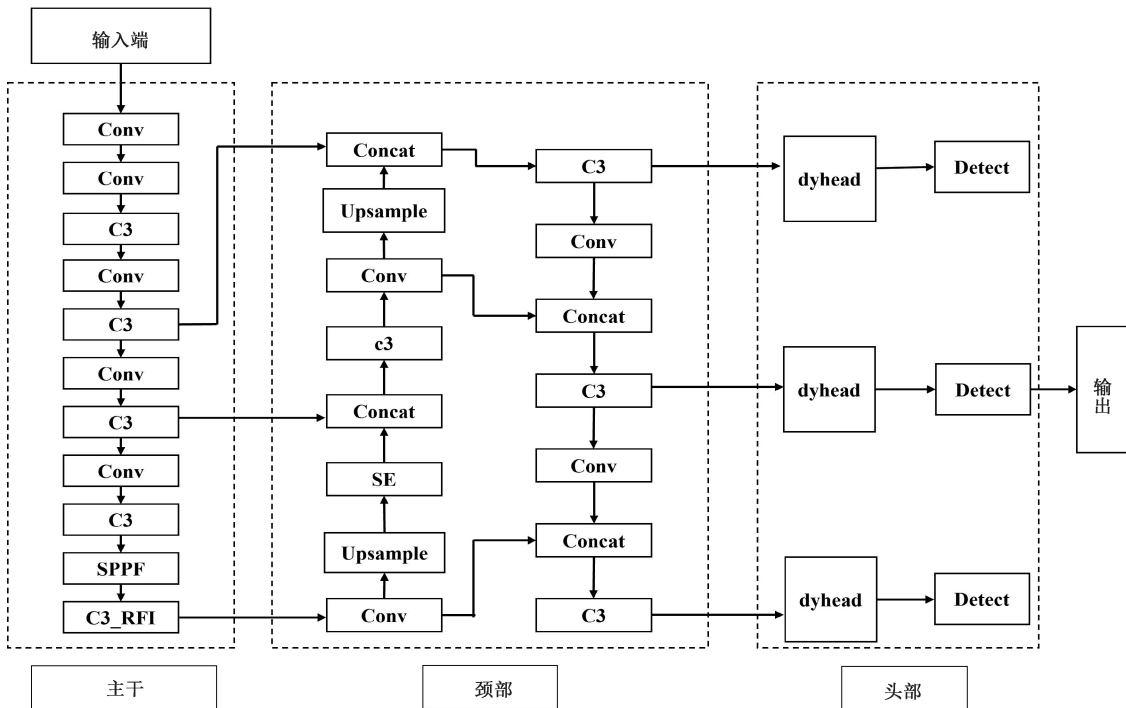


图 5 太阳能电池片表面缺陷检测模型结构

模型结构由主干、颈部、头部 3 部分组成。其中,主干以 CSPDarknet53 为骨干,通过特征映射和跨层次连接将电池片特征分解且合并,在减少计算成本的同时提高了特征提取的精度。颈部通常用于将不同层级的特征图结合起来,生成具有多尺度信息的特征图,以提高目标检测的准确率。头部用于分类和回归目标的位置,以及特征金字塔进行目标检测的部分。

2.2 检测模型优化

本文使用 YOLOv5s 通用目标检测器,应用在太阳能

电池片表面缺陷检测上,以解决电池片表面缺陷检测精度低、实时性差等问题。如图 6 所示,在主干中,用 P5 层的 RFI 模块代替瓶颈,融合多尺度特征。在颈部继续维持空间金字塔 SPPF (spatial pyramid pooling-fast, SPPF) 和路径聚合网络 PAN (path aggregation network, PAN) 的结构。此外,为了提高目标位置感知能力,将 P2 层集成到 PAN 中。在头部添加动态头部 dyhead,用于分类和回归目标的位置,增强模型小目标检测能力。具体来说,图 6(a)部分是检测器的主干,由跨阶段局部连接 CSP (cross stage

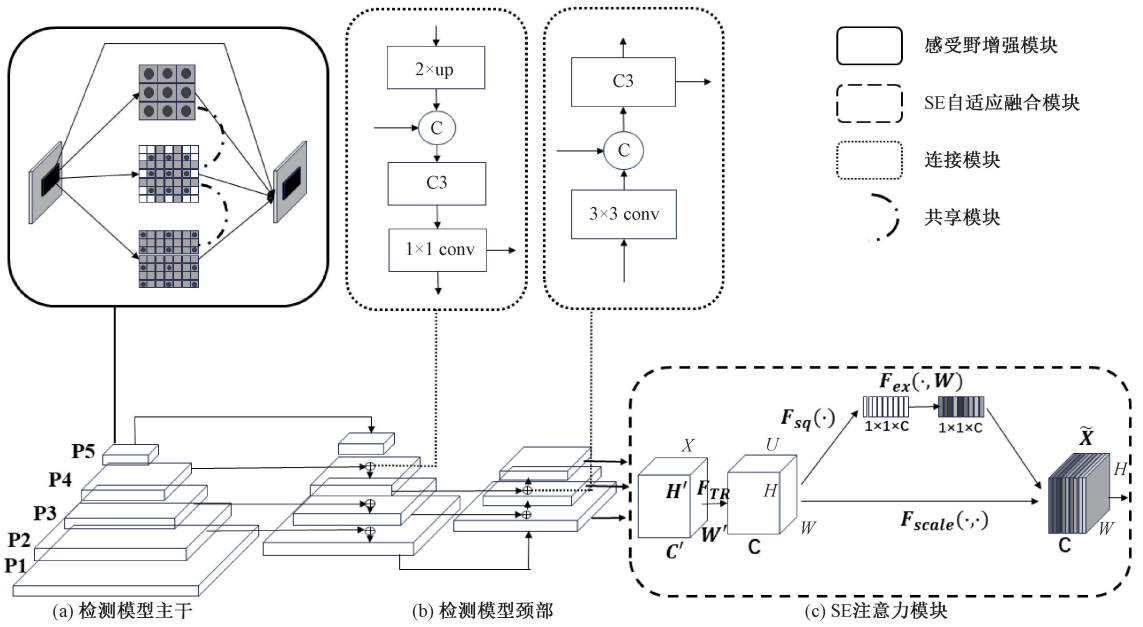


图 6 改进模型框架图

partial, CSP)块和跨阶段双向特征融合 CBS(cross-stage bi-directional feature fusion, CBS)块组成,用于提取输入图像的特征。在 P5 层中加入 RFI 模块,扩大有效感受场,增强多尺度融合能力。图 6 (b)部分称为颈部,由 SPPF 和 PAN 组成,将 P2 层的特征融合到颈层,来提高目标的定位能力。图 6 (c)部分引入 SE 注意力模块对输入特征图进行通道特征加强,以提高分类和定位精度。综合以上措施,可以显著提升太阳能电池片表面缺陷检测的准确度和实时性。

1) SE 自适应融合模块

本文使用 SE^[16]注意力模块,通过显式地建模其卷积特征通道之间的相互依赖性来提高网络产生的表示的质量。该注意力机制通过注意力模块的整体实现,可分为 2 步。

第 1 步:通道依赖性问题,为了解决利用通道依赖性问题,将每个学习到的过滤器都与接受域一起操作,并使用全局平均池来生成通道统计信息来实现通道依赖性的解决如式(13)所示,对应图 4 中的 F_{sq} 。

$$Z_c = F_{sq}(U_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j) \quad (13)$$

式中: Z_c 为 $Z \in R^c$ 通过空间维度 $H \times W$ 对 U 收缩生成的统计量, U_c 为输入 $U \in R^{H \times W \times c}$ 的特征映射, H, W, C 为特征的高度、宽度、通道数。

第 2 步:自适应再校正,为了利用第 1 步操作聚合的信息,选用简单的 S 型激活门控机制。首先,该机制比较灵活,可以学习通道之间的非线性相互作用。其次,该机制可以学习非互斥关系,确保多个通道被激活如式(14)所示,对应图 6 中的 $F_{ex}(\cdot, W)$ 。

$$S = F_{ex}(z, W) = \sigma(W_2 \delta(W_1 z)) \quad (14)$$

式中: W_1, W_2 分别用于降维和升维的两个全连接层权重, δ 为激活函数 relu 。

2) 归一化高斯沃瑟斯坦距离

YOLOv5s 自带的 IoU 损失函数对大物体的检测具有良好的效果,但对小物体的检测却不太理想。针对太阳能电池片表面缺陷的小物体缺陷,本文添加归一化沃瑟斯坦距离 $NWD^{[16]}$, 添加在 YOLOv5s 的定位 loss 中来弥补 IoU 损失对小目标检测的不足。首先,将边界框建模为二维高斯分布,然后通过它们对应的高斯分布来计算它们之间的相似性。

使用最优运输理论中的沃瑟斯坦距离来计算分布距离,对于两个二维高斯分布边界框 $A = (cx_a, cy_a, W_a, h_a)$ 和 $B = (cx_b, cy_b, W_b, h_b)$ 之间的高斯沃瑟斯坦距离定义如式(15)所示。

$$W_2^2(N_a, N_b) = \left\| \left(\left[cx_a, cy_a, \frac{W_a}{2}, \frac{h_a}{2} \right]^T, \left[cx_b, cy_b, \frac{W_b}{2}, \frac{h_b}{2} \right]^T \right) \right\|_2^2 \quad (15)$$

但 $W_2^2(N_a, N_b)$ 是距离变量,不能用作相似性度量(介于 0 和 1 之间的值作为 IoU)。因此,利用它的指数形式归一化,得到新的度量:称为归一化沃瑟斯坦距离 NWD 如式(16)所示。

$$NWD(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right) \quad (16)$$

式中: N_a, N_b 是边界框 $A = (cx_a, cy_a, W_a, h_a)$ 和 $B = (cx_b, cy_b, W_b, h_b)$ 的高斯分布, $W_2^2(N_a, N_b)$ 是最优运输理论中的沃瑟斯坦分布距离, C 是与数据集密切相关的常数。

3) 比例感知模型

因为接收野大小的不同意味着捕获远程依赖的能力

的不同,本文通过添加 RFI^[17] 模块以及扩展卷积来充分利用特征映射中接收野的优势。如图 7 所示,RFI 模块使用 4 个不同扩张卷积速率的分支来捕获多尺度信息和不同的依赖范围,从而充分利用每个样本,并减少参数数量,降低潜在过拟合的风险。RFI 模块可分为基于扩展卷积的多分支和采集加权层两部分,其中多分支部分使用固定的卷积核大小 3×3 ,增加了残差连接以防止训练过程中梯度爆炸和消失的问题。采集加权层用于收集来自不同分支的信息,并对每个分支的特征进行加权,以平衡不同分支的信息。

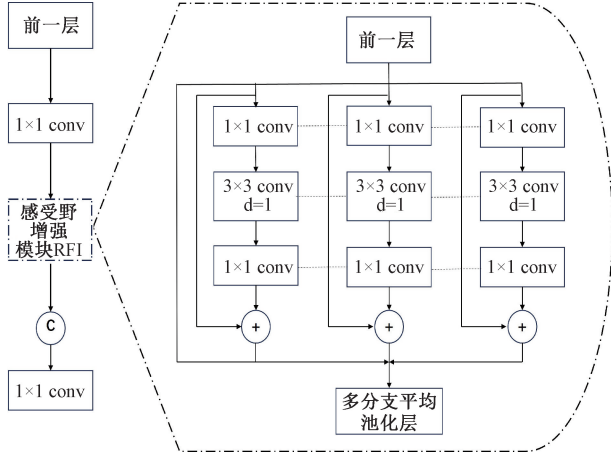


图 7 比例感知模块

4) 动态头部

近年来,针对目标检测器的改进主要集中在动态头部的研究上,包括尺度感知、空间感知和任务感知等方面。本文添加了动态头部 dyhead^[18] 框架,使各种特征级别的重要性自适应于输入。同时,通过空间感知将注意力应用于每个空间位置,并自适应地将多个特征级别聚合在一起以学习更具判别性的特征。为了进一步提高模型的复杂性和能力,本文将动态卷积(dynamic convolution, DConv)^[19] 添加到动态头部 dyhead 中。该动态卷积可以在不增加网络深度或宽度的情况下增加模型的复杂性,并通过注意力聚合多个卷积核来提高模型的识别能力。

在给定特征金字塔中 L 个不同级别的特征 $F_m = \{F_i\}_{i=1}^L$ 的连接,用上采样或下采样来调整特征的大小,使其接近中位数水平特征的规模。所以特征金字塔就可以表示为一个四维张量 $F \in \mathbf{R}^{L \times H \times W \times C}$,其中 L 表示金字塔的层数, H 、 W 、 C 分别表示中位层特征的高度、宽度和通道数。接着定义 $S = H \times W$ 于是将张量定义为 $F \in \mathbf{R}^{L \times S \times C}$ 。

给定张量 $F \in \mathbf{R}^{L \times S \times C}$,一般的自我注意的公式如式(17)所示。

$$W(F) = \pi(F) \cdot F \quad (17)$$

式中: $\pi(\cdot)$ 是一个注意函数,该注意力功能的解决方案通过全连接层来实现,但直接学习所有维度的注意函数对计

算能力的要求过高。因此,将注意力函数转换成 3 个连续的注意,每个注意只关注一个角度如式(18)所示。

$$W(F) = \pi_C(\pi_S(\pi_L(F) \cdot F) \cdot F) \cdot F \quad (18)$$

式中: $\pi_L(\cdot)$ 、 $\pi_S(\cdot)$ 和 $\pi_C(\cdot)$ 分别是作用于维度 L 、 S 、 C 上的 3 个不同的注意函数。 $\pi_L(\cdot)$ 是尺度感知关注,可以根据不同的尺度的语义重要性动态融合特征。 $\pi_S(\cdot)$ 是空间意识注意力,主要采用空间感知注意力模块来关注空间位置和特征层次之间一致判别区域。 $\pi_C(\cdot)$ 是任务感知注意力。

3 实验结果与分析

3.1 实验配置

本实验使用 Windows 11 操作系统,CPU 为 12 代 Intel(R) Core(TM) i7-12700H@2.30 GHz,内存为 16 GB,GPU 为 NVIDIA GeForce RTX 4060 Laptop,显存为 16 GB。软件环境包括 Anaconda3、Pycharm2022.3 以及深度学习库 PyTorch 1.7 和 CUDA 11.3。

本实验中,使用 Labellmg 工具对太阳能电池片表面缺陷数据集进行缺陷类型标记,生成相应的标签文件,其中包含了太阳能电池片的划痕、黑点、无电、叉隐、良品等 5 种缺陷类型。标注完成后,对模型训练路径等相关参数进行初始化,图像输入尺寸设定为 640×640 ,batch size 设为 4,权重衰减系数设为 0.0005,学习速率设为 0.001。

3.2 消融实验

本文在自己的太阳能电池片表面缺陷的数据集上对各个模块做了相关实验,评估各个模块对模型性能的影响。各个模块的对比分析如表 2 所示。

1) 自适应融合模块

本文添加了 SE 注意力机制,以加强模型特征层次的空间编码质量。其结果如表 2 第 2 行所示,准确率、召回率、mAP 分别提高了 2.9%、7.3%、2.2%。

2) 比例感知模型

本文先在 C3 部分添加了多元分支模块 DBB(diverse branch block,DBB)来丰富特征空间,但添加该模块后计算量负担过大。因此,在 C3 模块中添加了比例感知模块 RFI 来增加模型的感受野能力。其结果如表 2 第 4 行所示,准确率、召回率、mAP 分别提高了 8.1%、6.3%、4.6%。

3) 动态头部

本文添加动态头部 dyhead 框架通过在不同层级和通道之间动态调节注意力,从而提高模型的表征能力和检测精度。其结果如表 2 第 7 行所示,准确率、召回率、mAP 分别提高了 7.3%、8.7%、4.7%。

4) 动态卷积

本文在目标检测网络中引入了动态卷积 DConv 可以根据输入数据的不同特征来自适应地生成卷积核,从而提高模型的非线性表达能力。其结果如表 2 第 4 行所示,准确率、召回率、mAP 分别提高了 4.8%、7.0%、4.5%。

表 2 模块有效性消融实验

YOLOv5s	SE	C3_DBB	C3_RFI	Wiou(loss)	NWD(loss)	dyhead	DConv	准确率/%	召回率/%	mAP/%
✓								73.6	75.0	80.4
✓	✓							76.5	82.3	82.6
✓		✓						79.3	79.5	84.4
✓			✓					81.7	81.3	85.0
✓				✓				82.6	81.9	84.9
✓					✓			83.6	81.7	85.4
✓						✓		80.9	83.7	85.1
✓							✓	78.4	82.0	84.9
✓	✓		✓					80.6	82.4	85.0
✓	✓		✓		✓			84.8	81.6	85.8
✓	✓		✓		✓	✓		82.4	84.0	85.9
✓	✓		✓		✓	✓	✓	82.5	85.1	86.5

结果表明,添加的各个模块对太阳能电池片表面缺陷检测模型性能提升明显。由表 2 最后一行可知,改进后的 YOLOv5s 模型相比于 YOLOv5s,准确率、召回率、mAP 分别提高了 8.9%、10.1%、6.1%。

5)归一化沃瑟斯坦距离

本文先使用了基于 IoU 的动态非单调调频损耗 Wise-IoU(WIoU)来评价锚盒的质量。WIoU 可以专注于普通质量的锚盒,降低高质量锚盒的竞争力,也减少了低质量样例产生的有害梯度。但在微小物体的检测方面,WIoU 没有达到预期的效果。因此本文将 YOLOv5s 原始的 IoU 损失函数与归一化沃瑟斯坦距离 NWD 进行融合,通过一定的比例关系来提高小目标缺陷检测的性能。其结果如表 3 第 3 行所示,准确率、召回率、mAP 分别提高了 7.8%、8.0%、4.7%。

表 3 损失函数融合比例

CIoU	NWD	准确率/ %	召回率/ %	mAP/ %	Epochs
1	0	76.5	82.3	82.6	100
0	1	81.0	79.6	84.0	100
0.5	0.5	81.4	83.0	85.1	100
0.4	0.6	80.6	82.4	85.0	100
0.6	0.4	80.3	82.8	84.5	100

3.3 对比实验

为了进一步验证模型的检测性能,在保持训练平台配置信息不变、数据集相同的情况下,首先与 YOLOv5s、YOLOv7、YOLOv8 进行对比,接着与 Faster-RCNN、YOLOX-S 检测模型进行对比,最终结果如表 4 所示。

由表 4 可知,首先与 YOLOv5s、YOLOv7、YOLOv8 相比,改进的 YOLOv5s 的准确率分别提高了 11.9%、34.5%、2.8%,召回率分别提高了 10.1%、14.7%、4.8%,

表 4 模型对比实验

网络模型	准确率/ 召回率/		mAP/ %	参数量/ 计算量/	
	%	%		M	G
YOLOv5s	73.6	75.0	80.4	7.02	15.8
YOLOv7	51.0	70.7	58.8	37.22	105.2
YOLOv8	82.7	80.3	85.4	11.13	28.4
Faster-RCNN	75.0	71.2	77.4	138.2	33.2
YOLOX-S	73.4	79.7	79.7	43.85	19.3
改进 YOLOv5s	85.5	85.1	86.5	7.32	16.5

mAP 分别提高了 6.1%、27.7%、1.1%,然后与 Faster-RCNN、YOLOX-S 检测模型进行对比,改进后的模型准确率分别提高 10.5%、12.1%,召回率分别提高 13.9%、5.4%,mAP 分别提高 9.1%、6.8%。本文改进的模型,首先与 YOLOv7 和 YOLOv8 相比参数和计算量都大幅下降,可以很大程度满足生产实时性的要求,然后与 YOLOv5s 相比在略微增加模型参数量和计算量的情况下,可以大幅增加缺陷的检测精度,最后与 Faster-RCNN、YOLOX-S 相比参数量和计算量均大幅下降。通过与以上模型对比,充分说明改进后的模型对太阳能电池片表面缺陷检测的性能更强。

4 结 论

本文提出了一种基于虚拟相机的图像透视校正和改进的 YOLOv5s 模型,对太阳能电池片进行图像校正和表面缺陷检测。图像透视校正是一种对图像进行几何变换的技术,在计算机视觉和鱼虾识别至关重要。本文利用张正友标定获得的相机外参解出轴旋转角度和变换矩阵,构建虚拟相机,从而建立原图到虚拟相机正视图的透视映射关系并作用于电池片图像,从而实现太阳能电池片的快速透视校正。同时改进的 YOLOv5s 模型通过添加动态头部 dyhead、感受野增强模块 RFI、NWD Loss 损失函数、SE

注意力模块,提高了模型的特征提取和小缺陷目标的定位精度,加快了模型收敛速度,减少了过拟合的风险。通过基于虚拟相机的透视校正和改进 YOLOv5s 模型,太阳能电池片的图像校正速度提高 120 ms,校正效果更自然,减少传统校正图像方法在对图像校正时图像质量的损失。并在几乎不增加模型参数和计算量的情况下,太阳能电池片表面缺陷检测平均精度提高 6.1%,对太阳能电池片在工业生产中的质量保障有现实意义。

参考文献

- [1] MARIA P M, DEMETRIO P, GENNARO P. Values, concern, beliefs, and preference for solar energy: A comparative analysis of three European countries [J]. *Environmental Impact Assessment Review*, 2022, 93:106722.
- [2] BREYER C. Low-cost solar power enables a sustainable energy industry system[J]. *Proceedings of the National Academy of Sciences*, 2021, 118(49): e2116940118.
- [3] 周志锋,万旺根,王旭智. 基于 YOLO V3 框架改进的目标检测[J]. *电子测量技术*, 2020, 43(18): 102-106.
- [4] 周晓彦,王珂,李凌燕. 基于深度学习的目标检测算法综述[J]. *电子测量技术*, 2017, 40(11): 89-93.
- [5] 鲁东林,王淑青,鲁濠,等. 一种改进 FasterR-CNN 的太阳能电池片缺陷检测方法[J]. *激光杂志*, 2022, 43(3): 50-55.
- [6] 王宪保,李洁,姚明海,等. 基于深度学习的太阳能电池片表面缺陷检测方法[J]. *模式识别与人工智能*, 2014, 27(6): 517-523.
- [7] 张会林,唐政,马立新. 密集主干网络算法对太阳能电池的缺陷检测[J]. *激光与光电子学进展*, 2023, 60(14): 1412003.
- [8] 王淑青,朱文鑫,张子言,等. 基于改进 YOLOX-S 的太阳能电池片表面缺陷检测[J]. *激光杂志*, 1-6.
- [9] 杨元喜,徐天河. 不同坐标系综合变换法[J]. *武汉大学学报(信息科学版)*, 2001, 26(6): 509-513.
- [10] 罗晓晖,杜召彬. 基于双灭点的图像透视变换方法[J]. *计算机工程*, 2009, 35(15): 212-214.
- [11] KIM Z W. Robust lane detection and tracking in challenging scenarios [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2008, 9(1): 16-26.
- [12] CLARK P, MIRMEHDI M. Rectifying perspective views of text in 3D scenes using vanishing points[J]. *Pattern Recognition*, 2003, 36(11): 2673-2686.
- [13] DENG Z, ZHOU L. Detection and recognition of traffic planar objects using colorized laser scan and perspective distortion rectification [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 19(5): 1485-1495.
- [14] REDMON J, DIVVALA S K, GIRSHICK R B, et al. You only look once: Unified, real-time object detection[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016; 779-788.
- [15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018; 7132-7141.
- [16] WANG J, XU C, YANG W, et al. A normalized Gaussian Wasserstein distance for tiny object detection [EB/OL]. (2022-06-14) [2023-08-03]. <https://arxiv.org/abs/2110.13389>.
- [17] YU Z, HUANG H, CHEN W, et al. Yolo-facev2: A scale and occlusion aware face detector [EB/OL]. (2022-08-04) [2023-07-20]. <https://arxiv.org/abs/2008.02019>.
- [18] DAI X, CHEN Y, XIAO B, et al. Dynamic head: Unifying object detection heads with attentions[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021; 7373-7382.
- [19] CHEN Y, DAI X, LIU M, et al. Dynamic convolution: Attention over convolution kernels[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020; 11030-11039.

作者简介

朱栋,副教授,硕士生导师,主要研究方向为图像处理、计算电磁学(FDTD)、微纳光子学、数字信号处理等。

E-mail: zhudong@cczu.edu.cn

胡伟笑,硕士研究生,主要研究方向为图像处理、计算机视觉。

E-mail: s22060858008@smail.cczu.edu.cn

赵腾,硕士研究生,主要研究方向为图像校正、计算机视觉。

E-mail: 1353216629@qq.com