DOI:10.19652/j.cnki.femt.2305870

融合改进 YOLOv7 与 UNet 的编码点定位方法*

刘 升[1] 古丽巴哈尔·托乎提[1,2] 补生来[1] 买买提明·艾尼[1]
(1.新疆大学智能制造现代产业学院 乌鲁木齐 830047;
2.西安交通大学机械制造系统工程国家重点实验室 西安 710049)

摘 要:针对编码点的定位存在误检率高和精准度差的问题,提出了一种融合改进 YOLOv7 与 UNet 的圆形编码点定位方法。第1阶段使用改进的 YOLOv7 检测编码点的位置,改进的 YOLOv7 首先将 DCN-v2 可变形卷积引入 ELAN 模块,提升特征提取能力;其次把卷积块注意力模块(CBAM)机制嵌入骨干网络使网络更关注目标特征;然后使用 Focal-EIoU Loss 提高收敛速度;最后构建 OD-Cat 模块替换 ConCat 模块以提升网络检测精度。提取出每个圆形编码点的 ROI 后,第2阶段通过 UNet 分割出编码点的中心轮廓后,使用最小二乘法拟合出编码点的中心。实验结果表明,改进后的模型比原 YOLOv7 的精确率提高了 6.33%,平均精度均值(mAP)提升了 5.76%;提出的定位方法验证了在噪声、亮度不足或曝光等复杂环境下可以准确定位出编码点的中心椭圆轮廓,在实际工业视觉测量中具备鲁棒性。

关键词:编码点识别;深度学习;中心定位;YOLOv7;UNet;最小二乘法
中图分类号: TP 391.4 **文献标识码**:A **国家标准学科分类代码**: 520.6040

Fusion of improved YOLOv7 and UNet codepoint localization methods

Liu Sheng[1] Gulbahar Tohti[1,2] Bu Shenglai[1] Mamtimin Geni[1]
(1. College of Intelligent Manufacturing Modern Industry, Xinjiang University, Urumqi 830047, China;
2. State Key Laboratory for Manufacturing Systems Engineering, Xi′an Jiaotong University, Xi′an 710049, China)

Abstract: Aiming at the problems of high misdetection rate and poor accuracy in the localization of codepoints, a circular codepoint localization method fusing improved YOLOv7 and UNet is proposed. In the first stage, the improved YOLOv7 is used to detect the location of coding points. The improved YOLOv7 firstly introduces DCN-v2 deformable convolution into the ELAN module to improve the feature extraction ability. Secondly, the CBAM attention mechanism is embedded into the backbone network to make the network pay more attention to the target features. Then, Focal-EIoU loss is used to improve the convergence speed. Finally, OD-Cat is constructed to replace the ConCat module to improve the network detection accuracy. Module to replace the ConCat module to improve the network detection accuracy. After extracting the ROI of each circular coding point, the center contour of the coding point is segmented by UNet in the second stage, and then the center of the coding point is fitted using the least squares method. The experimental results show that the improved model improves the precision by 6.33% and the mean average precision (mAP) by 5.76% over the original YOLOv7. The proposed localisation method verifies that it can accurately locate the central ellipse contour of the coded point under complex environments such as noise, insufficient brightness or exposure, and is robust in practical industrial vision measurements.

Keywords: coded point recognition; deep learning; center localization; YOLOv7; UNet; least squares method

收稿日期:2023-12-21

* **基金项目**:国家自然科学基金地区基金(12162031)、西安交通大学机械制造系统工程国家重点实验室(sklms2022022)项目资助

0 引 言

在视觉测量领域中，圆形编码点因其结构稳定、抗噪性能好和中心定位精确的优势应用十分广泛。然而随着工业需求的精度要求上升，编码点的定位和识别的可靠性要求也在不断提高。同时视觉检测应用的条件也在复杂化，如在纷乱的场景和背景环境、极端光照或噪声影响下，要求检测方案需要较高的准确性和鲁棒性。目前，国内外学者提出的编码点定位和识别方案主要分为两种：第1种是通过设计编码点特征并通过图像处理进行检测；第2种是基于深度学习算法自主学习编码点特征进行检测。

编码点的第1种检测方案如下，金韬等[1]通过对传统的编码点添加定位符的方法可以增加标志点的数量和确定标志点的精确位置，但定位符的存在增大了编码点的尺寸，缺失灵活性且不容易布设；张小迪等[2]设计了一种基于同心圆约束原理的编码点，但其编码缺少起始点，若改变编码点位置则无法保证编码点的唯一性；刘慧洁等[3]设计了一种通过数字与字母的码值和中心十字定位的圆形编码点，一定程度上提高了识别定位精度。此类编码点设计方案基于中心圆的定位和环绕编码带的解码，充分利用了圆在几何中的稳定性和旋转不变性。然而在检测时过于依赖于稳定和单一的拍摄环境，在复杂场景和环境下识别精确率低，甚至还会错误定位。因此编码点的准确的识别和定位方法仍是需要深入研究的问题。

近年来，深度学习已经成为机器视觉及图像处理的一个重要领域，目标检测算法[4-6]的快速发展使编码点的第2种检测方案研究取得了极大进展。秦大辉等[7]通过将YOLOv3算法[8]的主干网络替换为Darknet-53并引入残差结构，提高了编码点的识别精度；陈明军等[9]引入了一种卷积神经网络(convolutional neural network，CNN)，解决了在三维视觉测量中识别模糊编码目标点的难题；崔海华等[10]提出一种基于深度学习图像超分的圆形靶标稳定检测方法，在拍摄倾角较大、位置较远的情况下实现了较高的靶标定位精度；陶聪等[11]设计了一种基于卷积神经网络以汉字作为编码特征的检测识别方法。由此可见，通过深度学习去识别编码点特征的方案在复杂环境影响下依然具备可靠性。然而目前的编码点研究更加专注于码值的创新和解码的准确率。针对工业测量要求逐渐增高的现状，更需要考虑编码点的定位精度问题。

通过目标检测网络可以识别并裁剪出编码点的位置，隔绝来自外界复杂环境的干扰。然而存在的亮度与噪声问题等都严重干扰着编码点的中心识别。为了更加精准的定位编码点中心，继续引入了语义分割中的显著性目标检测。该模型可以通过网络学习生成目标对象的掩膜，进一步提高目标检测的精准度，常见的显著性目标检测模型有UNet[12]、UNet＋与DeepLabV3＋等。通过此类网络模型对编码点的每个像素进行分类，可以准确分割出编码点的中心圆。

综上所述，为了增强目标检测算法对编码点的识别精确性和鲁棒性，本文提出融合改进YOLOv7[13]与UNet的编码点识别算法。首先改进YOLOv7主干网络的卷积块提升特征提取能力，更加有效的获取编码点的特征，引入可变形卷积增强空间不变性并扩大采样范围。其次嵌入了结合通道维度和空间维度的注意力机制获取丰富特征信息，抑制无关信息的干扰，使网络更加关注图像中目标的特征。然后引入更专注于高质量锚框的损失函数增高回归精度，增加对目标框宽度和高度的损失计算，减少了目标框与锚框的宽度差和高度差，从而提高回归精度和收敛速度。最后通过UNet分割出编码点中心圆，避免复杂环境中极端光照与噪声的干扰，使用最小二乘法拟合出精确的编码点中心椭圆轮廓。

1 编码点定位流程

编码点的设计在解码方式与编码容量上各有不同，然而在中心识别上均采用了中心圆定位方案，如图1(a)～(d)所示。而传统识别方式依赖稳定的拍摄环境，流程如图2(a)所示，该方法往往受限于复杂环境中的边缘检测干扰和阈值处理精度的问题。针对编码点的识别和定位精度问题，本文基于深度学习的视觉算法模型提出了融合两个网络模型的算法去定位编码点中心：首先通过改进后的YOLOv7从复杂环境出提取出每一个编码点的感兴趣区域(region of interest，ROI)，然后将每个编码点送入UNet网络分割出每个编码点的中心位置，最后使用基于最小二乘法的椭圆识别得到精准的像素坐标，实现流程如图2(b)所示。深度学习视觉算法精准度和泛化性的优势，往往可以使模型快速且精准的识别出准确位置。

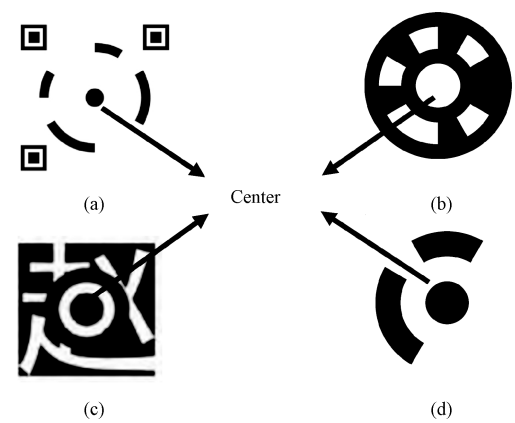

图1 各类编码点

Fig. 1 Various coded points

2 改进YOLOv7模型

为了定位圆形编码点的中心位置，提出了一种改进的YOLOv7算法模型。改进后的YOLOv7网络模型如图3所示。首先本文引入DCN-v2可变形卷积替换主干的CBS常规卷积和E-ELAN的3×3卷积，增强主干网络的

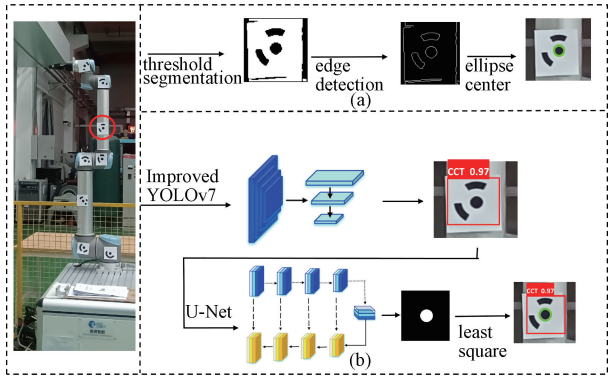

图 2　网络流程

Fig. 2　Net process

特征提取能力；其次在主干网络嵌入 CBAM 注意力机制使其更加关注目标的特征信息；然后构建 OD-Cat 模块令网络在采样流程可以捕获丰富的网络上下层信息，提升检测能力；最后使用 Focal-EIoU Loss 损失函数提高回归精度和检测速度。

2.1　改进模块 DC-ELAN

ELAN 模块通过控制梯度路径的长短关系，使网络能够学习更多特征，然而大量使用的 CBS 卷积层会导致该模块特征提取效果不佳。为了提升主干网络的特征学习能力，使用了 DCNv2 卷积模块替换 3×3 常规 Conv 卷积。该卷积对常规引入了偏移量，可针对卷积操作区域进行尺寸和比例的变化，使卷积区域变形为不规则的卷积，并寻找有用的信息区域。

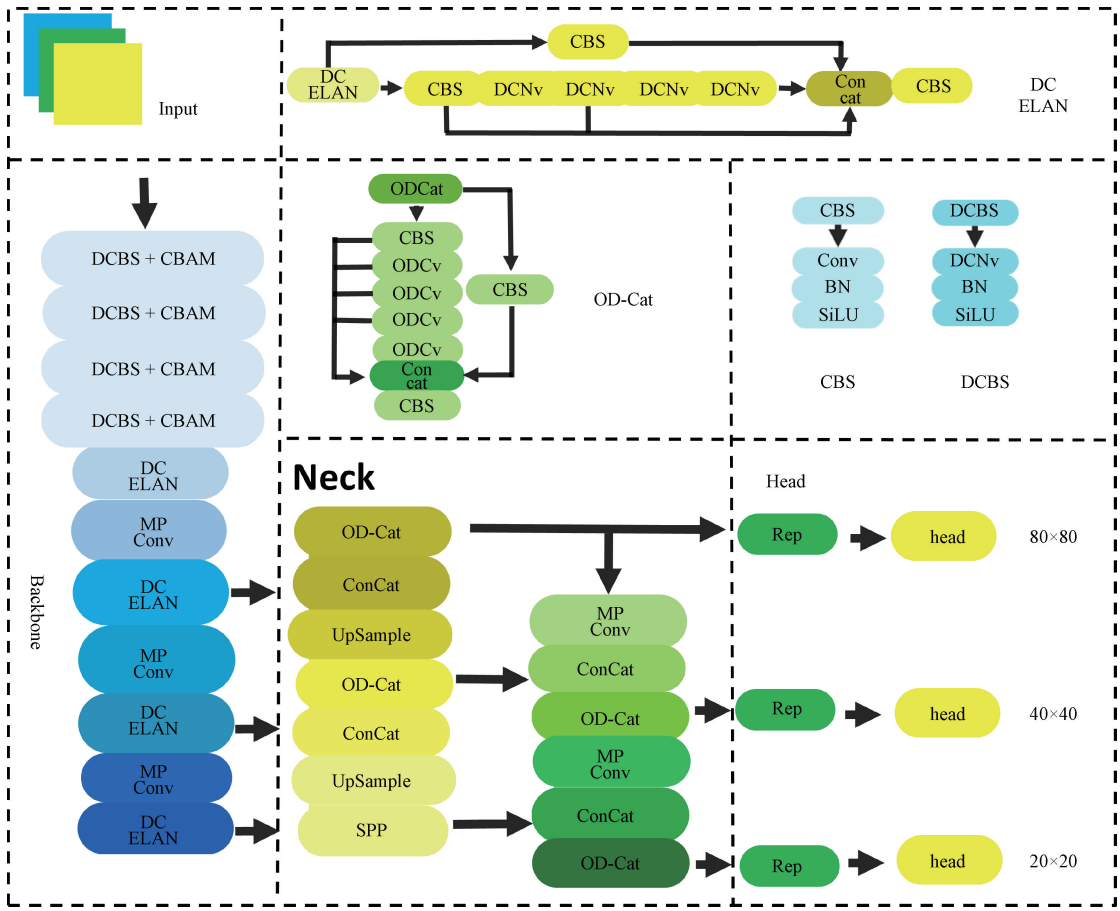

图 3　改进 YOLOv7 模型

Fig. 3　Improved YOLOv7 model

DCNv2 可变形卷积增强空间不变性并扩大采样范围，可以通过堆积调制过的 DCN 块来使得模型的表现力强于普通的 CNN 的堆积[14-15]，具体采样方式如图 4 所示。

可变性卷积操作并没有改变卷积的计算操作，而是在卷积操作的作用区域上，加入了一个可学习的权重参数 Δm_k，如式(1)所示。

$$Y_{(p_0)}=\sum_{k=1}^{k} w_k \cdot X(p+p_k+\Delta p_k)\cdot \Delta m_k \tag{1}$$

式中：$p+p_k+\Delta p_k$ 表示待卷积的新采样位置；X 和 Y 分别表示输入和输出的特征图；Δm_k 是通过卷积进行学习的参数，评估引入区域是否为感兴趣的区域，防止对错误区域进行识别，同时更加确定有效信息的精准提取。

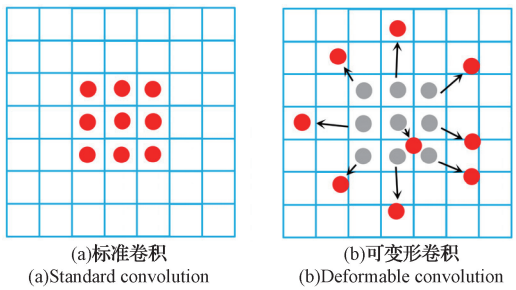

图 4　DCNv2 卷积

Fig. 4　DCNv2 convolution

2.2　卷积块注意力模块(CBAM)

为了提高对编码点的特征提取效果，将 CBAM 注意力机制[16]嵌入 YOLOv7 主干网络中。CBAM 注意力机制结合了通道维度和空间维度，可以抑制无关信息的干扰，使网络更加关注图像中目标的特征，结构如图 5 所示。

其中通道注意力模块(channel attention module，CAM)对输入的特征图进行最大池化运算和平均池化运算，将特征图变为 $C\times1\times1$ 的大小，经过多层感知机(multi-layer perceptron，MLP)将通道数缩小 r 倍，然后将压缩后的通道数扩张为原始的通道数，并根据元素对两个

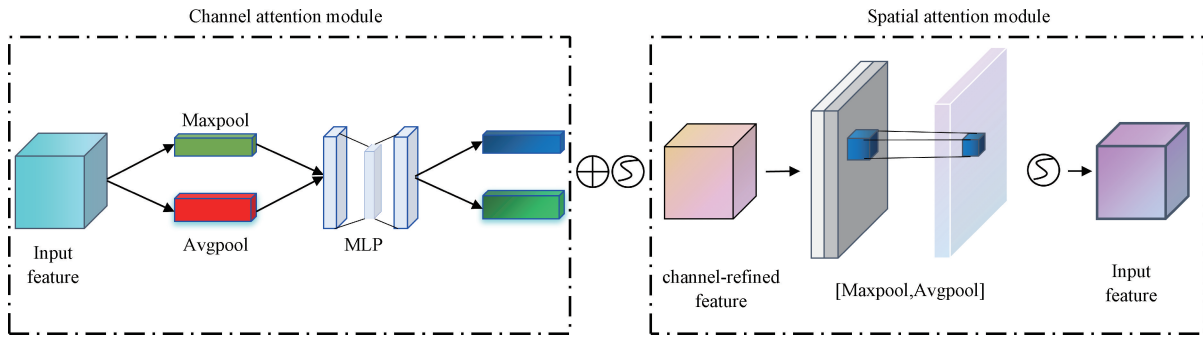

图 5　CBAM 结构

Fig. 5　CBAM structure

输出求和，然后送入 Sigmoid 激活函数输出结果。空间注意模块(spatial attention module，SAM)将通道注意模块的输出作为输入，经过最大池化运算和平均池化运算后得到两个 $1\times H\times W$ 的特征图，并将两个特征图拼接在一起，经过卷积和 Sigmoid 激活函数运算后得到整个 CBAM 模块的输出。

CBAM 注意力机制的通道注意力模块与空间注意力模块计算公式如下：

$$M_cF=\sigma\{MLP[AP(F)]+MLP[MP(F)]\} \tag{2}$$

$$M_sF=\sigma\{f^{7\times7}[AP(F)];MP(F)]\} \tag{3}$$

式中：σ 为激活函数 Sigmoid；$f^{7\times7}$ 为 7×7 的卷积；AP 为平均池化 AvgPool；MP 为最大池化 MaxPool。

2.3　损失函数改进

为解决预测框与真实图像数据不重叠产生的梯度消失问题，YOLOv7 引入了 Focal-EIoU Loss[17]，更精准的计算预测框与真实框的边界定位损失。

Focal-EIoU Loss 由 EIoU Loss 和 Focal Loss 整合而成。Focal Loss 的引入是从梯度的角度考虑问题，从而区分高质量和低质量的锚框，使回归过程更专注于高质量的锚框。

Focal-EIoU Loss 的公式如下：

$$L_{Focal-EIOU}=IOU^{\gamma}L_{EIOU} \tag{4}$$

式中：γ 是控制异常值抑制程度的参数。

EIoU Loss 由 3 部分组成，重叠损失、中心距离损失以及宽度和高度损失，前两部分的损失函数延续了 YOLOv7 中 CIoU 的计算方法，但是宽度和高度损失减少了目标框与锚框的宽度差和高度差，从而加快了收敛速度，提高了回归精度。其中惩罚项计算公式如下：

$$L_{EIoU}=L_{IoU}+L_{dis}+L_{asp}=1-\frac{(A\cap B)}{(A\cup B)}+\frac{\rho^2(b,b^{gt})}{c^2}+\frac{\rho^2(w,w^{gt})}{C_w^2}+\frac{\rho^2(h,h^{gt})}{C_h^2} \tag{5}$$

式中：A 为真实标记框的面积；B 为预测边缘框的面积；C_w 和 C_h 是覆盖两个真实标记框的最小外接矩形的宽度和高度；$\rho^2(b,b^{gt})$ 是预测框中心点到真实标记框的中心点的欧氏距离；w^{gt}、h^{gt}、w 和 h 分别是预测边界框与真实标记框的宽和高。

2.4　改进模块 OD-Cat

YOLOv7 的 Neck 部分的采用上采样与下采样专注于特征强化，其中的 Cat-Conv 在采样后实现特征融合效果。使用全维度动态卷积 OD-Conv[18]替换 Cat-Conv 中的 1×1 和 3×3 的 Conv。动态卷积模块同时考虑了空域、输入通道、输出通道等维度上的动态性，通过并行策略采用多维注意力机制沿着核空间的 4 个维度分别学习注意力。这 4 种类型的注意力是互补的，通过渐进式方案对卷积沿位置、通道、滤波器以及核空间等维度乘以不同的注意力将使得卷积操作对于输入存在各个维度的差异性，提供更好的性能以捕获丰富上下文信息，可以大幅提升卷积的特征提取能力，动态卷积 ODConv 如图 6 所示。

动态卷积通过通道全局的均值池化操作(GAP)，将输

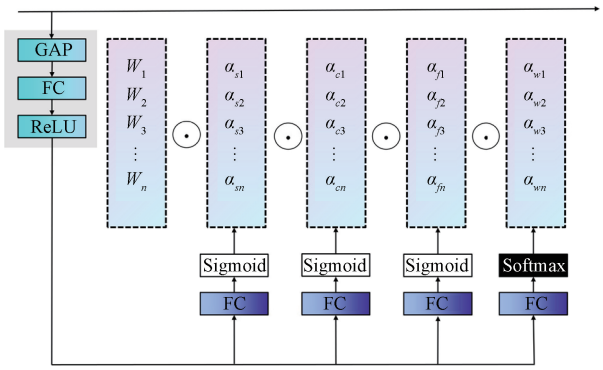

图 6 ODConv 结构

Fig. 6 ODConv structure diagram

入压缩为一个特征向量；随后，经过全连接层(FC)和激活函数(ReLU)；再次经过 FC 层将压缩后的特征向量映射到低维空间；最后使数据通过 4 个分支分别生成 α_{si}、α_{ci}、α_{fi} 和 α_{wi} 的 Softmax 或 Sigmoid 函数。

ODConv 使用 SE 类型的 4 个注意力模块，具体如下式所示：

$$y = (\alpha_{w1} \odot \alpha_{f1} \odot \alpha_{c1} \odot \alpha_{s1} \odot W_1 + \cdots + \alpha_{wn} \odot \alpha_{fn} \odot \alpha_{cn} \odot \alpha_{sn} \odot W_n) * x \tag{6}$$

式中：α_{wi} 表示为卷积核的数量注意力；α_{ci} 表示对输入通道数的注意力；α_{fi} 表示对输出通道数的注意力；α_{si} 表示为卷积核的空间尺寸注意力。

3 UNet 模型与中心定位

改进 YOLOv7 模型提取到圆形编码点 ROI 区域后，若通过传统的图像处理方式：阈值分割、边缘检测等进行目标识别时会存在复杂背景环境的干扰问题，将导致编码点的识别和定位精度存在很大误差。因此本文采用显著性目标检测 UNet 网络避免环境的干扰，从而分割出编码点的轮廓，以保留编码点中心的亚像素信息。

UNet 网络结构如图 7 所示，主要分为下采样阶段(encoder)和上采样阶段(decoder)。下采样阶段使用两个

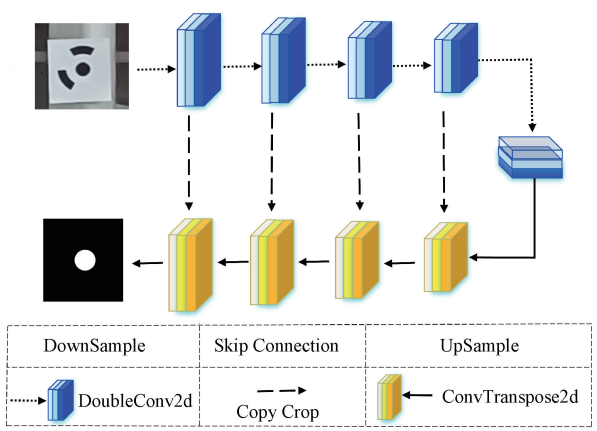

图 7 UNet 结构模型

Fig. 7 UNet structure model

卷积层专注于提取编码点的特征，上采样阶段则是结合转置卷积和跳跃连接得到高分辨率的特征图，实现对每一个像素点的类别划分，得到编码点的精确轮廓。

编码中心点的椭圆通过对 UNet 网络生成的掩膜图像进行最小二乘法拟合。椭圆方程式的一般表达式如下：

$$F_{(\alpha,X)} = \alpha \cdot X = ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{7}$$

其中：

$$\alpha = (a, b, c, d, e, f) \tag{8}$$

$$X = (x^2, xy, y^2, x, y, 1) \tag{9}$$

根据最小二乘法，通过最小化误差的平方和，拟合出椭圆方程如下：

$$F_{(a,b,c,d,e,f)} = \sum_{i=1}^{n} (ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i + f)^2 \tag{10}$$

将椭圆的参数计算后，可以进一步根据椭圆方程计算出位置参数与形状参数如下：

$$\begin{cases} x_0 = \dfrac{2bc - ad}{a^2 - 4b} \\ y_0 = \dfrac{2d - ac}{a^2 - 4b} \end{cases} \tag{11}$$

$$\begin{aligned} A &= \sqrt{\frac{2(acd - bc^2 - d^2 + 4be - a^2e)}{(a^2 - 4b)(b + 1 - \sqrt{a^2 + (1-b)^2})}} \\ B &= \sqrt{\frac{2(acd - bc^2 - d^2 + 4be - a^2e)}{(a^2 - 4b)(b + 1 + \sqrt{a^2 + (1-b)^2})}} \\ \theta &= \tan^{-1}\sqrt{\frac{A^2 - B^2b}{A^2b - B^2}} \end{aligned} \tag{12}$$

式中：(x_0, y_0)是椭圆位置参数，中心坐标；A、B、θ 是形状参数，椭圆长轴、短轴与倾角。

4 实验结果和分析

4.1 实验数据集

本文使用的数据集为自制数据集，检测目标为 Schneider 编码点。为了验证该实验的可行性，如图 8 所示，首先收集了不同倾角与不同光照下的编码点图像，共拍摄采集了 654 张包含大量编码点的图像数据；其次通过翻转、添加不同类型的噪声、裁剪和旋转等图像预处理方式增强数据集；最后通过增强效果将数据集扩大到包含 3 270 张照片(4 608×3 456)并将其作为改进 YOLOv7 模型的数据集。该数据集划分比例为 8∶1∶1 的训练集、测试集和验证集。使用 LabelImg 软件对 3 250 图像中的编码点进行单目标的标注。

通过改进 YOLOv7 识别出自制数据集中的编码点，裁剪为含单一编码点的图片，随机选择 2 500 张图像作为显著性目标检测模型 UNet 的数据集，然后通过 Labelme 软件标注出编码中心圆点作为分割的目标进行训练，如图 8(g)与(h)所示，其中图 8(h)中红色编码中心点是识别目标。

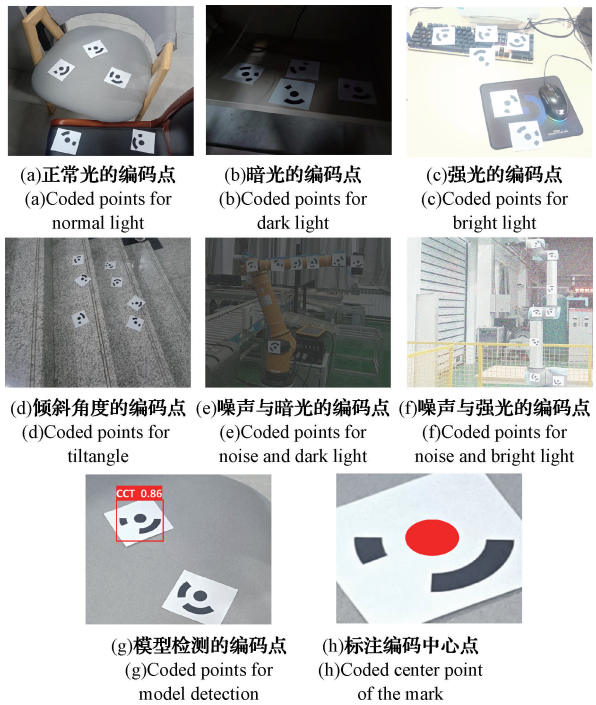

图 8　复杂环境下编码点数据集

Fig. 8　Coded point datasets in complex environments

4.2　网络模型训练

1)实验环境

本文实验使用的平台是 PyTorch1.10 深度学习框架，计算机操作系统为 Windows10，Cuda 版本为 cuda11.3，Python 版本为 3.8，CPU 为 14 核 Intel(R) Core(TM) i5-12400F，GPU 为 NVIDIA RTX 4060(8 GB)。

2)训练细节

训练时使用官方的 YOLOv7.pt 预训练权重，数据集使用自制编码点数据集，标签选择 YOLO 格式，输入图像尺寸为 640×640，迭代次数 Epoch 设置为 150 轮，采用 Adam 优化器，初始学习率为 0.01，Batchsize 根据性能稳定性设置为 4。

UNet 训练数据集中随机选择 10%作为验证集，随时检测网络的训练情况。优化训练采用 Adam 优化器，Batchsize 为 8，输入图像为 224×224，迭代次数为 200，初始学习率为 0.001。

4.3　评价指标

实验用精确率(precision，P)、召回率(recall，R)、平均精度均值(mAP)对改进模块的效果进行评判。

精确率是指预测为正的正样本所占的比例，可以衡量模型错误的检测程度。召回率指的是预测结果为正的真实样本百分比，可以衡量模型的遗漏检测程度。精确率和召回率的公式如下：

$$P = \frac{TP}{TP + FP} \tag{13}$$

$$R = \frac{TP}{TP + FN} \tag{14}$$

式中：TP 为正确检测到的目标个数；FP 为错误检测到的目标个数；FN 为没有检测到的正确目标个数。

mAP@0.5 是指把 IoU 设置为 0.5 时，计算每个类别的所有图像的平均精度(AP)后对所有类别求平均值，用于衡量模型准确率随召回率变化的趋势。mAP 公式如下：

$$mAP = \frac{\sum_{i=1}^{C} AP_i}{C} \tag{15}$$

式中：C 为类别数。

4.4　消融实验

为了验证改进 YOLOv7 网络算法的有效性和可靠性，针对本文提出的 4 处改进通过逐一增加改进模块的方法设计了消融实验，结果如表 1 所示。消融实验通过“√”表示使用该策略，“×”表示未使用该策略。

表 1　改进消融实验

Table 1　Improved ablation experiment

组别	CBAM	DC-ELEN	Focal-EIoU Loss	ODConv	Precision/%	Recall/%	mAP@0.5/%
A	×	×	×	×	89.45	84.26	88.86
B	√	×	×	×	91.53	85.79	91.89
C	√	√	×	×	92.31	88.67	89.67
D	√	√	√	×	93.29	87.11	93.21
E	√	√	√	√	95.78	90.34	94.62

实验 A 使用原 YOLOv7 网络模型在数据集上的实验结果，其中精确率为 89.45%，召回率为 90.26%，平均精度均值为 90.86%。

实验 B 添加了 CBAM 注意力机制后抑制无关信息的干扰，使网络更加关注图像中目标的特征。精确率提高了 2.08%，平均精度均值提高了 1.03%。

实验 C 在实验 B 的基础上，将主干网络的 ELAN 模块替换为 DC-ELAN 模块，CBS 卷积替换为包含可变形卷积的 DCNv2 卷积块，召回率提高了 4.41%。精确率上升了 2.86%。

实验 D 基于实验 C 使用 Focal-EIoU Loss 损失函数，平均精度均值提高了 2.35%，精确率提高了 3.84%。

实验 E 是本文提出的改进 YOLOv7 网络模型，相比实验 D 基础上在 YOLOv7 的 Neck 部分处的 OD-Cat 中，将卷积替换为动态卷积 ODConv 模块，模型的精确率提高了 6.33%，召回率提高了 6.08%，平均精度均值提高了 5.76%。

4.5 对比实验

为了验证本文改进算法的有效性，使用自制数据集和实验设备，对本文网络与其他网络进行了测试。这些网络包括 YOLOv3、YOLOv5s、SSD、Faster R-CNN、YOLOv8 和 Center-Net。实验结果如表 2 所示，本文改进算法效果优于经典的 one-stage 与 two-stage 算法，且本文提出的模型平均精度均值，精确率，召回率都是最高。

表 2 各模型对比试验

Table 2 Comparative experiments of each model (%)

Model	Precision	Recall	mAP@0.5
YOLOv3[19]	82.64	82.61	80.90
Center-Net[20]	86.79	80.98	83.64
YOLOv5s	91.78	87.16	87.07
SSD[21]	80.95	70.27	79.64
Faster R-CNN[22]	79.96	71.95	78.92
YOLOv8	89.79	85.63	88.75
本文	95.78	90.34	94.62

4.6 编码点定位测试

为验证本文提出的编码点定位方法的实用性与鲁棒性，在自制数据集中选择编码进行视觉算法检测。随机选择图 9 中标注的 7 个编码点作为测试对象。

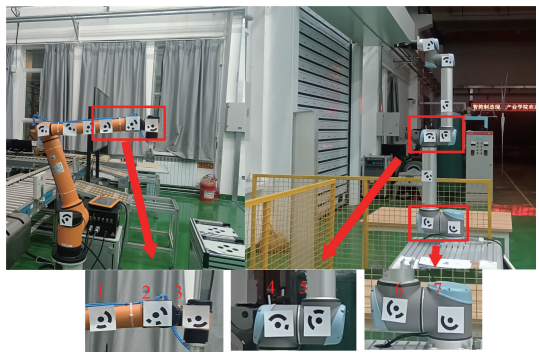

图 9 编码点识别测试

Fig. 9 Coded point detection test

在复杂背景下(噪声、强光照与弱光照)识别和定位编码点的中心，测试效果如图 10 所示。

CCT 传统算法是通过最小二乘法去拟合圆形编码点中心的椭圆。在弱光照与噪声的影响下，误识率极高，甚至将编码带识别为编码中心点；在强光照与噪声的影响下，噪声的像素信息更大程度的干扰了最小二乘法对于编码点中心椭圆轮廓的判断；在正常光照与噪声的影响下，传统算法收到的干扰较小，基本可以识别出编码中心位置，但仍存在较大误差。

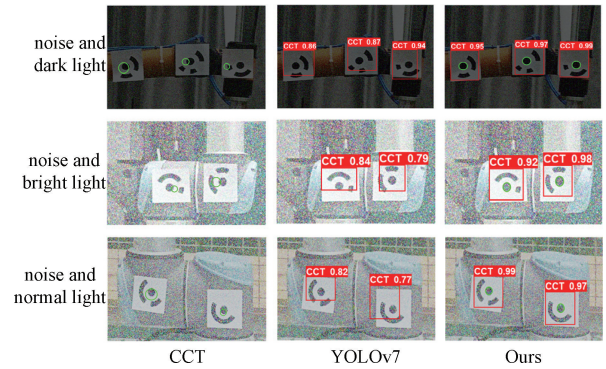

图 10 算法验证效果

Fig. 10 Algorithm test result

YOLOv7 算法在弱光照与噪声的影响下，检测出了包含编码中心点的大概位置且识别准确率达到 85%以上，然而编码区域没有正确识别；在强光照和噪声以及正常光照和噪声的影响下，识别效果均低于 85%，远远达不到实际应用需求。

本文提出的融合改进 YOLOv7 与 UNet 的中心定位方案可以准确定位出每个编码点的区域且定位出编码中心位置。由图 10 可以看出，复杂背景干扰下，中心轮廓的识别精度效果相比较 CCT 传统算法明显提升。而改进 YOLOv7 的测试效果，弱光照、强光照与噪声等复杂环境影响下编码点的中心与编码带均在识别范围内，平均准确率达到 95.78%以上。该方案呈现的精确率与鲁棒性满足使用要求。

5 结 论

本文提出了一种融合改进 YOLOv7 与 UNet 的编码点识别算法，第 1 阶段以 YOLOv7 为基础模型，使用 DC-ELEN 模块、CBAM 注意力机制、OD-Cat 模块，Focal-EIoU Loss 损失函数改进并优化原网络模型。在自制数据集上实现优于原 YOLOv7 模型和其他的目标检测算法，模型的精确率提高了 6.33%，召回率提高了 6.08%，平均精度均值提高了 5.76%。在复杂环境中识别出编码点的位置，同时便于编码点后续的解码和中心定位识别。第 2 阶段采用 UNet 精准分割出编码点的中心位置，并通过最小二乘椭圆拟合去定位中心圆轮廓。本文提出的算法在实验中展现出远高于传统编码识别方案的精度效果，表现出极佳的鲁棒性，更加适用于工业测量。本文采用传统的 Schneider 编码点进行了测试和研究，该方案同时适用于不同类型的编码点定位，也可以应用于实际物体的精确定位。下一步将构建轻量化与实时性的编码点识别框架平台，以便于将其应用于真实的工业近景摄影检测任务。

参考文献

[1] 金滔，董秀成. 一种环状编码标记点的设计及解码算法[J]. 计算机应用研究，2019，36(4)：1217-1221.
JIN T，DONG X CH. Designing and decoding algorithm of circular coded target[J]. Application Research of Computers，2019，36(4)：1217-1221.

[2] 张小迪，崔海华，程筱胜，等. 一种基于同心圆的圆形编码标志设计与检测[J]. 机械制造与自动化，2020，49(2)：133-136,157.
ZHANG X D，CUI H H，CHENG X SH，et al. Design and detection of circular marker based on concentric circles[J]. Machine Building & Automation，2020，49(2)：133-136,157.

[3] 刘慧洁，买买提明·艾尼，古丽巴哈尔·托乎提，等. 一种合作编码标志的设计与识别定位[J]. 激光与光电子学进展，2021，58(12)：199-207.
LIU H J，GENI M，TOHTI G，et al. Design and identification of cooperative coded targets[J]. Laser & Optoelectronics Progress，2021，58(12)：199-207.

[4] 侯学良，单腾飞，薛靖国. 深度学习的目标检测典型算法及其应用现状分析[J]. 国外电子测量技术，2022，41(6)：165-174.
HOU X L，SHAN T F，XUE J G. Analysis of typical target detection algorithm based on deep learning and its application status[J]. Foreign Electronic Measurement Technology，2022，41(6)：165-174.

[5] 范先友，过峰，俞建峰，等. 基于改进YOLOv7的液晶面板电极缺陷视觉检测技术研究[J]. 电子测量与仪器学报，2023，37(9)：225-233.
FAN X Y，GUO F，YU J F，et al. Research on visual detection technology for liquid crystal panel electrode defect by improved YOLOv7[J]. Journal of Electronic Measurement and Instrumentation，2023，37(9)：225-233.

[6] 娄文，郭杜杜，张杰，等. 基于YOLOv7的驾驶人使用手机与抽烟行为识别方法[J]. 电子测量技术，2023，46(21)：123-131.
LOU W，GUO D D，ZHANG J，et al. Identification method of mobile phone use and smoking behavior ofdrivers based on YOLOv7[J]. Electronic MeasurementTechnology，2023，46(21)：123-131.

[7] 秦大辉，程栋，苏铭著，等. 基于目标检测网络的摄影测量编码点定位[J]. 激光与光电子学进展，2021，58(4)：182-189.
QIN D H，CHENG D，SU M ZH，et al. Photogrammetric coded point localization based on target detection network[J]. Laser & Optoelectronics Progress，2021,58(4)：182-189.

[8] REDMON J，DIVVALA S，GIRSHICK R，et al. You only look once：Unified，real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR)，2016：779-788.

[9] 陈明军，周含策，张丽艳. 基于卷积神经网络的运动模糊编码点识别[J]. 计算机辅助设计与图形学学报，2017,29(10)：1844-1852.
CHEN M J，ZHOU H C，ZHANG Y L. Recognition of motion blurred coded targets based on convolutional neural network[J]. Journal of Computer-Aided Design & Computer Graphics，2017，29(10)：1844-1852.

[10] 崔海华，徐振龙，杨亚鹏，等. 一种基于深度学习图像超分的圆形靶标稳定检测方法[J]. 中国机械工程，2021,32(23)：2861-2867.
CUI H H，XU ZH L，YANG Y P，et al. A circular target stability detection method based on deep learning image super-resolution[J]. China Mechanical Engineering，2021，32(23)：2861-2867.

[11] 陶聪，施云，张丽艳. 基于卷积神经网络的汉字编码标记点检测识别[J]. 仪器仪表学报，2019，40(8)：191-200.
TAO C，SHI Y，ZHANG Y L. Detection and recognition of Chinese character coded marks based on convolutional neural network[J]. Chinese Journal of Scientific Instrument，2019，40(8)：191-200.

[12] RONNEBERGER O，FISCHER P，BROX T. UNet：Convolutional networks for biomedical image segmentation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention(MICCAI). 2015：234-241.

[13] WANG C Y，BOCHKOVSKIY A，LIAO H. YOLOv7：Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. Conference on Computer Vision and Pattern Recognition (CVPR). IEEE，2023：7464-7475.

[14] DAI J F，QI H Z，XIONG Y W，et al. Deformable convolutional networks[C]. IEEE International Conference on Computer Vision (ICCV). 2017：764-773.

[15] ZHU X Z，HU H，LIN S，et al. Deformable ConvNets V2：More deformable，better results[C]. 2019 Conference on Computer Vision and Pattern Recognition (CVPR). IEEE，2019：9300-9308.

[16] WOO S，PARK J，LEE J Y，et al. CBAM：Convolutional block attention module[C]. Proceedings of the European Conference on Computer

Vision (ECCV), 2018: 3-19.

[17] ZHANG Y, REN W, ZHANG Z, et al. Focal and efficient IoU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.

[18] LI C, ZHOU A J, YAO A B, et al. Omni-dimensional dynamic convolution [C]. 10th International Conferenceon Learning Representations, 2022.

[19] EDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. Computer Science, 2018, arXiv. 1804. 02767.

[20] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points[J]. Computer Science, 2019, arXiv:1904. 07850.

[21] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Singleshot multibox detector[C]. Proceedings of the European Conference on Computer Vision, 2016: 21-37.

[22] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

作 者 简 介

刘升,硕士研究生,主要研究方向为计算机视觉与图像处理。

E-mail:2065686217@qq.com

古丽巴哈尔·托乎提(通信作者),副教授,主要研究方向为计算机视觉与图像处理。

E-mail:gulbahart@163.com

补生来,硕士研究生,主要研究方向为计算机视觉与图像处理。

E-mail:1540971800@qq.com

买买提明·艾尼,教授,主要研究方向为计算机视觉与图像处理。

E-mail:mgheni@263.net