doi：10.11823/j.issn.1674-5795.2025.03.10

基于卷积神经网络的压力仪表OCR系统

王晶星，陈诗琳，李一鸣，王丽*，石伟，刘芳
（中国航空工业集团公司北京长城计量测试技术研究所，北京 100095）

摘　要：为了解决压力仪表传统人工抄表效率低、易出错且存在安全风险以及基于传感器和三维视觉的自动抄表技术适应性不足的问题，结合计算机视觉与人工智能技术，构建了融合数据采集、实时监控和数据分析功能的计量系统。通过改进快速区域卷积神经网络（Fast Region-Convolutional Neural Network，Fast R-CNN）算法，引入数据增强和轻量化特征提取网络，提高复杂环境下的仪表定位准确度；在DeepLabv3+模型中融合通道注意力机制与空间注意力机制，结合混合损失函数，提升字符分割效率。实验表明：改进算法在复杂工业环境中的仪表表盘定位平均准确度达84%，字符分割平均交并比为78.6%，单次计量耗时较人工抄表缩短85%，验证了系统的高效性与强适应性。本研究为工业设备智能监测提供了可扩展的技术框架，对推动计量数字化、智能化具有实践价值。

关键词：光学字符识别；设备现场计量；深度学习；仪表定位；字符分割
中图分类号：TB93；TP316；TH7　　**文献标志码：**A　　**文章编号：**1674-5795（2025）03-0111-12

OCR system for pressure instruments based on convolutional neural network

WANG Jingxing, CHEN Shilin, LI Yiming, WANG Li*, SHI Wei, LIU Fang
(AVIC Changcheng Institute of Metrology & Measurement, Beijing 100095, China)

Abstract: To address the inefficiency, error-proneness, and safety risks associated with traditional manual meter reading for pressure instruments, as well as the limited adaptability of automated meter-reading technologies based on sensors and 3D vision, this study integrates computer vision and artificial intelligence technologies to develop a metering system that combines data acquisition, real-time monitoring, and data analysis. By improving the fast region-convolutional neural network (Fast R-CNN) algorithm through data augmentation and a lightweight feature extraction network, the system optimizes instrument positioning accuracy in complex environments. Additionally, the DeepLabv3+ model is enhanced by incorporating channel attention and spatial attention mechanisms, along with a hybrid loss function, to improve character segmentation efficiency. Experimental results demonstrate that the improved algorithm achieves an average positioning accuracy of 84% for instrument dial positioning and a mean Intersection over union of 78.6% for character segmentation in challenging industrial environments. Furthermore, the system reduces the time required for a single measurement by 85% compared to manual reading, confirming its high efficiency and strong adaptability. This research provides a scalable technical framework for intelligent monitoring of industrial equipment, offering the practical value for advancing digital and intelligent metering.

Key words: OCR; equipment on-site metering; deep learning; instrument positioning; character segmentation

收稿日期：2025-02-14；**修回日期：**2025-03-11
基金项目：国家重点研发计划资助项目（2022YFB3207000）
引用格式：王晶星，陈诗琳，李一鸣，等. 基于卷积神经网络的压力仪表OCR系统[J]. 计测技术，2025，45(3)：111-122.
Citation: WANG J X, CHEN S L, LI Y M, et al. OCR system for pressure instruments based on convolutional neural network[J]. Metrology & Measurement Technology, 2025, 45(3): 111-122.

0 引言

在工业生产领域，压力参数对生产过程的安全性、稳定性和产品质量具有决定性影响。然而，目前其数据处理方式主要依赖人工读数和记录，这种传统压力仪表现场计量[1]方式效率低下，且容易受到操作人员主观因素的影响；此外，人工记录的数据难以实现实时传输和分析，无法满足现代工业生产对数据实时性和智能化管理的需求。现有的自动抄表技术虽然在一定程度上解决了上述问题，但仍存在局限性。例如，基于传感器的自动抄表技术虽然能够实现精准实时的数据采集，但需要对仪表进行改造，且成本较高；基于三维视觉的自动抄表技术虽然可消除图像畸变，但硬件成本高昂，且计算复杂度较高。相比之下，基于光学字符识别(Optical Character Recognition, OCR)[2]的自动抄表技术无需对仪表进行硬件改造，且成本较低，在复杂工业现场具有广阔的应用前景。

近年来，随着信息技术的飞速发展，OCR技术应运而生，并在诸多领域得到了广泛应用。OCR技术能够高效、准确地将图像中的文字和图像信息转换为计算机可识别的数据。通过将OCR技术应用于压力仪表的现场计量，研究人员[3]成功实现了压力数据的自动采集、快速识别和实时传输，有效弥补了传统现场计量方式的不足。在该领域，康耐视公司[4]开发的视觉检测系统采用了先进的图像处理算法。这些系统通过对图像预处理、特征提取与匹配，能在复杂背景下精准定位压力仪表的表盘并识别字符。然而这些系统的硬件成本较高，限制了其在一些场景中的广泛应用。商汤科技[5]的OCR系统通过大量压力仪表图像的训练，实现了对压力数据的高准确率识别，但该系统在实际应用中易受到环境因素的干扰，一定程度上影响了其稳定性和可靠性。

本文针对复杂计量现场的仪表定位与识别问题，提出了一种综合优化方案，通过优化图像预处理流程，并结合数据增强和优化网络架构等技术手段，有效减轻了光线不均、污渍沾染、背景干扰等恶劣条件对图像质量的负面影响；此外，本文提出了一种改进的分割算法，深入分析压力仪表字符的分布规律与特殊性，有效攻克了字符粘连、部分残缺等识别难题，确保了字符完整提取。在此基础上，本文将仪表定位、字符分割与OCR识别技术深度融合，构建了一套一体化的智能计量系统。该系统优化了各模块间的数据交互与协同工作流程，实现了高效、稳定的数据处理。通过这种创新的系统架构，成功满足了长期、稳定、高准确度计量的需求，为工业计量现场的智能化升级提供了有力支持。

1 OCR技术原理

1.1 OCR技术概述

OCR技术旨在将图像中的信息转化为机器能够理解与处理的编码文本。其核心原理是对图像中的字符进行特征提取和模式识别，借助一系列复杂算法实现从图像到文本的高效转换。以压力仪表识别为例，OCR技术流程如图1所示，具体实现过程如下：首先，利用摄像头、扫描仪等设备获取压力仪表表盘的图像；在图像预处理阶段[6]，针对采集图像可能出现的噪声、光照不均、倾斜等问题，运用滤波、灰度变换、倾斜校正等算法进行优化处理；接下来，在仪表盘定位阶段，通过特征提取、描述及匹配等步骤，精准定位仪表盘的位置；在字符分割阶段，依据压力仪表字符的排列规律、间隔特点等，将连续的字符序列分割为单个字符；随后，采用基于模板匹配、深度学习的字符识别算法，将分割后的字符与预定义的字符库或模型学习到的字符特征进行比对，从而确定字符类别；最终，输出准确、规范的压力仪表读数文本。

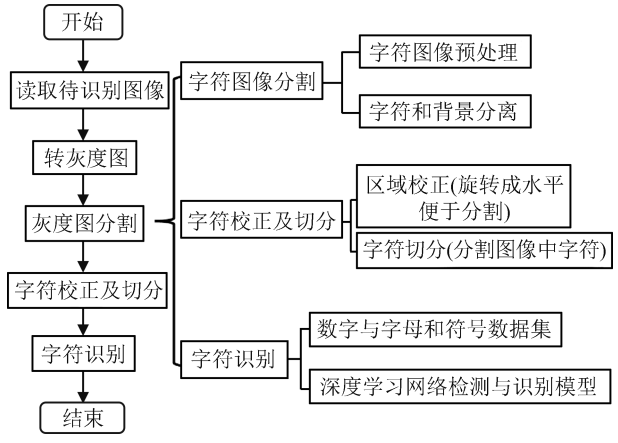


图1 OCR技术流程图

Fig.1 Flowchart of OCR technology

1.2　关键技术与算法

图像预处理是提升OCR识别准确率的基础。灰度化处理能将彩色图像转换为灰度图，有效简化数据量，同时突出字符轮廓；针对噪声干扰问题，中值滤波能够有效去除椒盐噪声，同时保持字符边缘的清晰度；对于光照不均问题，可使用自适应直方图均衡化方法，该方法能够根据图像局部区域的灰度分布动态调整对比度。

目标检测算法[7]能够在包含大量干扰信息的图像中，准确识别出压力仪表的位置和边界框。例如，在1张布满各种机械设备的图像中，该算法可以迅速锁定压力仪表所在的区域，将其从复杂的背景中分离出来，从而为后续OCR处理提供明确的目标范围。压力仪表的安装姿态多种多样，包括水平安装、垂直安装、倾斜安装。同时，在不同的拍摄距离下，仪表在图像中的尺度也会有显著差异。基于CNN[8]的目标检测模型，如Fast R-CNN、YOLO系列等，能够学习并识别压力仪表的特征，不管是面对不同的安装姿态(水平、垂直、倾斜)，还是不同的光照条件(强光、弱光、逆光)，都能实现快速且准确的目标定位。图2展示了基于CNN的目标检测的核心原理。

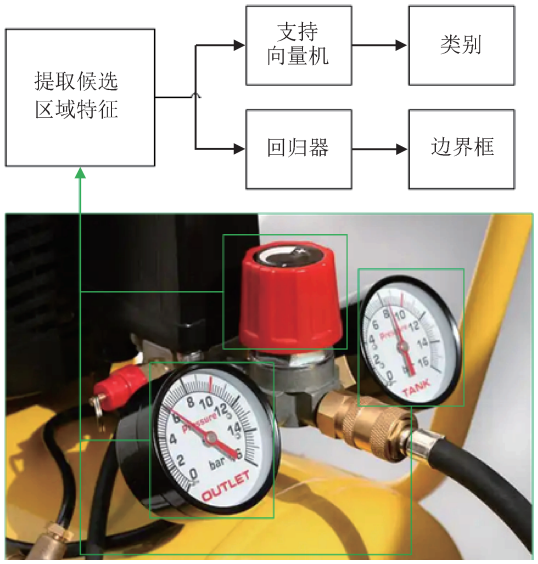


图2　基于CNN的目标检测模型原理

Fig.2　The principle of object detection models based on CNN

字符分割算法[9]是提升OCR识别准确度的关键环节。投影法依据字符在水平或垂直方向上的灰度投影变化，通过寻找波谷位置来确定字符边界，适用于字符间距均匀、排列规整的压力仪表；连通域分析法利用图像中像素的连通区域和几何特征筛选目标区域，能有效应对字符粘连的情况；基于CNN的深度学习分割算法通过大量样本学习字符形状、结构等特征，自动提取分割线，可对复杂字体、异形字符进行分割。

字符识别算法[10]决定OCR最终识别准确度。模板匹配算法将预先存储的字符模板与分割后的字符图像逐一比对，通过计算相似度来选取最匹配的模板对应的字符，在字符、字体、字号固定的专用压力仪表识别中，这种方法具有识别速度快、准确率高的优点；支持向量机(Support Vector Machine，SVM)[11]作为一种经典机器学习算法，通过将字符特征映射到高维空间，寻找最优分类超平面，能够在训练样本较少的情况下实现精准的字符分类，尤其适用于工业现场特定仪表字符集的识别；在深度学习领域，CNN凭借其多层卷积和池化层，能够自动学习字符的深层次特征，具有很强的鲁棒性，能适应多种字体以及模糊、变形字符。例如，在高温高压环境下，压力仪表的字符可能因表盘老化或污渍沾染而变得模糊不清，但CNN模型仍可对其准确识别。

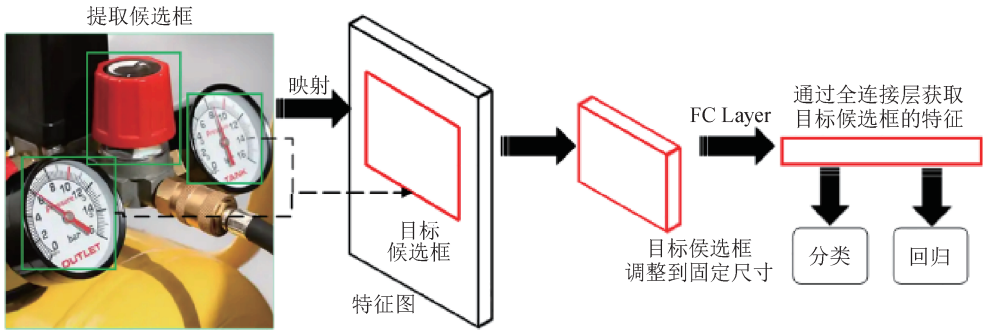
通过对这些关键技术与算法的深入剖析与合理选用，本文构建了一套基于OCR技术的压力仪表现场计量系统。

2　压力仪表现场计量中的仪表定位

在现场计量环境中，对压力仪表进行准确定位是后续识别的基础。在背景复杂、光照变化及遮挡等复杂环境下，传统目标检测方法难以实现对压力仪表的准确定位。为了解决这一问题，本文主要从增强引入数据和改进特征提取网络两个方面对Fast R-CNN算法[12](原理如图3所示)进行改进。

2.1　数据增强

数据增强[13]是一种利用计算机视觉中的多种图像处理技术对图像进行随机的几何变换和颜色调整的方法，它通过学习并模拟实际应用中可能遇到的各种不同光照强度、物体姿态和遮挡情况来扩充数据集。这一技术使得模型在训练过程中能够接触到更丰富的数据样本，从而提高模型的



注：FC Layer为全连接层(Fully Connected Layer)。

图3 Fast R-CNN原理图

Fig.3 Schematic diagram of fast R-CNN

泛化能力、增强其在复杂环境下的适应性，并有效减少过拟合风险。

数据增强过程作为数据预处理的一部分，在模型训练的输入阶段进行。具体流程如下：

1）从原始数据集中读取图像和对应的标签信息；

2）对图像依次进行亮度调整、对比度调整、旋转和缩放等随机变换操作，具体步骤如下：

① 亮度调整：在训练阶段，对每张图像的亮度进行随机调整，调整后的像素值I_{new}为

$$I_{new} = I_{old} \cdot (1 + \alpha) \tag{1}$$

式中：I_{old}为原始图像的像素值；α为$[-0.2, 0.2]$范围内的随机值，当α为正数时，图像亮度增加，当α为负数时，图像亮度降低。通过改变α可模拟出不同光照强度下的图像效果。

RGB(Red，Green，Blue)是一种加色模型，通过红(Red，R)、绿(Green，G)、蓝(Blue，B)三个颜色通道的数值组合表示色彩。例如，对于图像中的每个像素点(x, y)，其RGB值分别为r_{old}、g_{old}、b_{old}。按照亮度调整公式计算得到新的RGB值(r_{new}，g_{new}，b_{new})如式(2)所示。

$$\begin{cases} r_{new} = r_{old} \cdot (1 + \alpha) \\ g_{new} = g_{old} \cdot (1 + \alpha) \\ b_{new} = b_{old} \cdot (1 + \alpha) \end{cases} \tag{2}$$

将计算得到的新RGB值作为调整亮度后的像素值，遍历图像中的所有像素点，即可完成整幅图像的亮度调整。

② 对比度调整：基于图像的直方图对图像的对比度进行调整。首先计算图像的平均像素值$\bar{I}$，然后得出调整后的像素值I_{new}为

$$I_{new} = \beta \cdot (I_{old} - \bar{I}) + \bar{I} \tag{3}$$

式中：β为$[0.7, 1.3]$范围内的随机值，当$\beta > 1$时，图像对比度增强，当$\beta < 1$时，图像对比度减弱。通过这种方式，可以让模型适应不同对比度的图像。

③ 旋转：以图像中心为旋转中心，使用双线性插值法对图像进行旋转。旋转角度θ在$[-15°, 15°]$范围内随机选取。旋转矩阵$\boldsymbol{R}$为

$$\boldsymbol{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{4}$$

通过该矩阵对图像像素坐标进行变换。双线性插值法可以在旋转过程中对新像素值进行平滑计算，以减少图像失真。

④ 缩放：按照一定比例对图像进行缩放。随机生成缩放因子s在$[0.8, 1.2]$之间，将图像的宽和高分别乘以s，可以得到新的图像尺寸。使用双线性插值法填充缩放后得到的新图像的像素值，以保证图像的连续性和清晰度。

3）将变换后的图像和更新后的标签信息输入Fast R-CNN模型中进行训练，该模型可以确保在每次训练迭代时都能接收到不同的训练样本，从而增强模型的泛化能力。

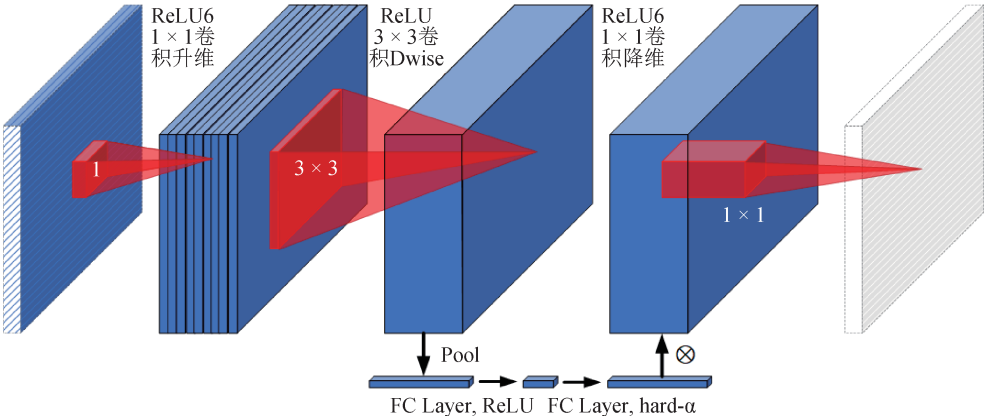
2.2 改进特征提取网络

Fast R-CNN中默认使用的VGG16网络[14]结构相对复杂且计算量较大，包含大量的卷积层和全连接层。VGG16在处理压力仪表这类小目标时容易丢失一些细节信息，导致特征提取不够准确。而MobileNetV3[15]采用了轻量级的深度可分离卷积技术，将传统的卷积操作分解为深度卷积和逐点卷积两个步骤，并引入线性瓶颈结构。通过在卷

积层之间引入线性激活函数，避免了非线性激活函数带来的信息损失，从而显著减少了参数数量和计算量。这两项技术手段的联合使用，不仅能减少计算量，还能通过引入注意力机制的Squeeze-and-Excitation结构，更高效地提取如压力仪表这类小目标的特征。

将Fast R-CNN模型中的VGG16特征提取网络直接替换为MobileNetV3网络，并在替换过程中确保MobileNetV3网络的输入尺寸与原Fast R-CNN模型对输入图像的要求一致，以便于调整后续分类和回归层与MobileNetV3输出特征的连接，从而使新的特征提取网络能够与Fast R-CNN的其他部分协同工作。MobileNetV3网络(原理如图4所示)的最后一层特征图作为感兴趣区域池化层的输入，经过感兴趣区域池化操作后，这些特征图被转换为固定大小的特征向量，随后输入到全连接层中进行分类和回归。

MobileNetV3网络由一系列的倒残差结构和Squeeze-and-Excitation(SE)模块组成，通过多个阶段逐步提取特征，最终得到具有代表性的特征图。每个倒残差结构由以下部分组成：首先是一个1×1的逐点卷积，用于升维；接着是一个3×3的特征提取深度卷积；最后通过一个1×1的逐点卷积完成降维操作。SE模块则对特征图进行全局平均池化，然后通过全连接层学习通道间的注意力权重，将这些权重应用到原始特征图上。



注：ReLU为激活函数；ReLU6为ReLU的变体；Dwise为深度可分离卷积；Pool为池化操作；FC Layer为全连接层(Fully Connected Layer)；hard-α为用于生成通道注意力权重的函数。

图4　MobileNetV3原理图

Fig.4　Schematic diagram of MobileNetV3

例如，一个包含C_{in}个输入通道和C_{out}个输出通道的倒残差结构计算过程如下：

① 升维阶段：输入特征图的通道数可通过1×1逐点卷积从C_{in}提升到C_{out}。计算量为$C_{in}\times C_{out}\times 1\times 1\times H\times W$，其中$H$和$W$分别为特征图的高和宽；

② 特征提取阶段：每个通道独立进行3×3深度卷积操作。计算量为$C_{out}\times 3\times 3\times H'\times W'$，其中$H'$和$W'$分别为卷积后特征图的高和宽；

③ 降维阶段：通道数通过1×1逐点卷积从C_{out}降回到C_{in}。计算量为$C_{out}\times C_{in}\times 1\times 1\times H'\times W'$。

2.3　实验验证

本次实验收集了1 000张压力仪表图像，这些图像涵盖了复杂的环境条件，包括不同的光照条件、背景干扰及遮挡情况。为了更好地进行模型训练与评估，数据集按照7:2:1的比例划分为训练集、验证集和测试集。

本文实验采用TensorFlow框架实现了算法改进。在训练过程中，初始学习率被设定为0.001，并采用Adam优化器进行参数更新，训练过程共进行100个epoch，每个epoch都会对训练集进行一次完整的遍历。此外，实验在训练过程中使用了早停策略，即当验证集上的损失函数在连续10个epoch内没有出现下降时就停止训练，以此有效避免过拟合现象发生。

本文实验采用平均准确度(Average Precision，AP)、召回率(Recall)、每秒浮点运算次数(FLOPs)和参数量(Params)作为评价指标，对模型进行了综合评估。

实验将改进后的Fast R-CNN算法与原始Fast R-CNN、YOLOv8和EfficientDet等目标检测算法进行对比分析，具体对比效果如表1所示。与原始Fast R-CNN相比，改进算法在测试集上的AP提升了9%，召回率也从0.7提升到了0.8，并且具有更小的模型复杂度。

表1 仪表定位算法效果对比

Tab.1 Effect comparison of instrument location algorithms

模型	AP/%	Recall	FLOPs/G	Params/M
原始Fast R-CNN	75	0.70	180	138
YOLOv8	80	0.82	62.6	11.1
EfficientDet	77	0.73	15.3	8.5
改进后的Fast R-CNN	84	0.80	90	45

改进算法在处理光照变化较大的图像时，定位准确率提高了15%。此外，对于遮挡率超过一半的压力仪表图像，改进算法将其定位准确率从原有的30%提升到了50%，显著提高了复杂环境下压力仪表定位性能的精准度。与其他目标检测算法相比，改进后的Fast R-CNN在准确度方面表现优异，但其FLOPs(90 G)和参数量(45 M)较高。因此，该算法更适合应用于资源充足且对准确度要求较高的压力仪表计量现场。复杂环境下的算法效果对比情况如图5所示。

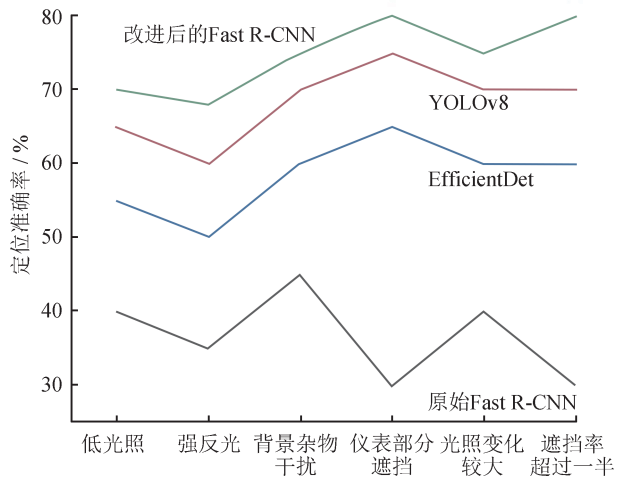


图5 复杂环境下的算法效果对比

Fig.5 Comparison of algorithm effects in complex environments

3 压力仪表元素分割算法

为了实现后续的压力仪表读数识别，通常采用传统算法对压力仪表的各个元素(如指针、刻度盘、数字显示等)进精确分割。然而，传统的图像分割算法在面对实际场景中复杂背景、光照变化以及仪表姿态等多样性问题时，难以满足高准确度分割的需求。

DeepLabv3+[16]是一种基于编码器-解码器架构的语义分割模型，它通常采用预训练的骨干网络(如ResNet、Xception等)作为编码器部分，用于提取图像的高级语义特征。同时，该模型利用空洞空间金字塔池化(Atrous Spatial Pyramid Pooling，ASPP)模块捕获不同尺度的上下文信息。解码器部分则负责将编码器输出的低分辨力特征图进行上采样并恢复至原始图像大小，并通过与浅层特征融合来恢复空间细节信息。

3.1 引入注意力机制

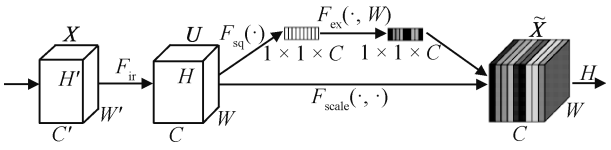
1）通道注意力机制

注意力机制能够让模型在处理特征时，自动关注到对目标识别更重要的区域。SE模块[17]通过学习通道间的依赖关系为不同通道的特征分配不同的权重。这使得模型在面对遮挡情况时能够更专注于未被遮挡部分的有效特征，从而提高对遮挡目标的定位能力。

SE模块可以自适应地调整每个通道的特征响应，增强重要特征的表达，抑制无关特征的干扰。在MobileNetV3的每个瓶颈结构之后添加SE模块，接收来自前一层的特征图，对其进行处理后输出带有注意力权重的特征图，再将其输入到后续网络层。通过这种方式，模型可以在特征提取的过程中不断地对特征进行自动调整和优化，更好地处理遮挡问题。

SE模块主要由挤压、激励和重标定三个部分组成。挤压部分通过全局平均池化将特征图压缩为一个1×1的向量，以获得每个通道的全局信息。激励部分可以调整两个全连接层学习通道间的注意力权重，第1个全连接层将通道数压缩为原来的1/r（r为压缩比，通常取16），并使用ReLU激活函数进行非线性变换；第2个全连接层将通道数恢复到原来的数量，并使用Sigmoid激活函数将权重

值映射到[0, 1]区间。重标定部分将学习到的权重与原始特征图相乘，得到带有注意力权重的特征图。SE模块原理如图6所示。



注：$\boldsymbol{X}$为输入的特征图，是一个三维张量，其高度为H'，宽度为W'，通道数为C'；F_{ir}为深度可分离卷积；U为经过F_{ir}操作后的特征图，其高度为H、宽度为W、通道数为C；$F_{sq}(\cdot)$为Squeeze操作函数，其作用是将特征图$\boldsymbol{U}$压缩为$1\times1\times C$的向量；$F_{ex}(\cdot, W)$为Excitation操作函数，是一个全连接层操作，W为全连接层的权重参数；$F_{scale}(\cdot,\cdot)$为Scale操作函数；$\widetilde{\boldsymbol{X}}$为SE模块的输出特征图。

图6　SE模块原理图

Fig.6　Schematic Diagram of SE Module

假设输入的特征图尺寸为$H\times W\times C$，SE模块的计算过程如图7所示。

① 挤压部分：对输入的特征图进行全局平均池化操作，将每个通道的二维特征图压缩为一个标量值，从而获取每个通道的全局信息，计算量为$H\times W\times C$，最终得到一个$1\times1\times c$的向量。

② 激励部分：首先通过第1个全连接层，其权重矩阵为$\boldsymbol{W}_1$，大小为$C\times(C/r)$，计算量为$C\times(C/r)$；接着通过ReLU激活函数；再通过第2个全连接层，其权重矩阵为$\boldsymbol{W}_2$，大小为$C\times(C/r)$，计算量为$C\times(C/r)$，最终得到通道注意力权重向量。

③ 重标定部分：将生成的通道权重与原始特征图逐通道相乘，对每个通道的特征进行加权处理，以突出重要通道的特征信息，计算量为$H\times W\times C$，最终得到重标定后的特征图。

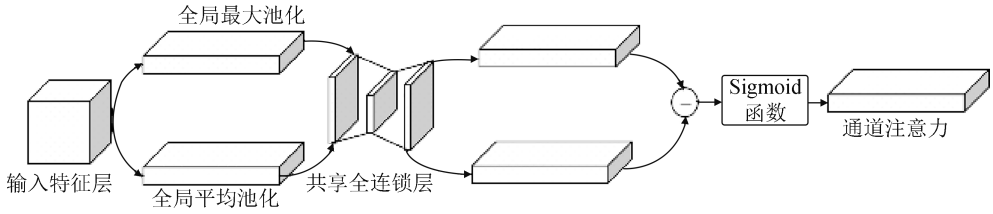


图7　SE模块计算过程

Fig.7　Calculation process of the SE module

2）空间注意力机制

在解码器将编码器特征与浅层特征融合之后，引入空间注意力模块(Spatial Attention Module, SAM)[18]。SAM模块的工作流程如图8所示。

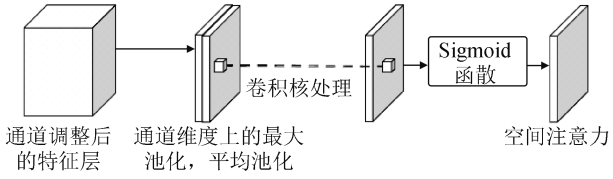


图8　SAM模块的工作流程图

Fig.8　Workflow diagram of the SAM module

① 对输入的特征图分别进行通道维度的最大池化和平均池化操作，从而得到两个不同的特征图。

② 将这两个特征图在通道维度上进行拼接，然后通过一个7×7的卷积层进行特征融合，生成一个表示空间注意力的特征图。

③ 使用Sigmoid函数将该特征图的值映射到[0, 1]区间，得到空间注意力权重图。

④将空间注意力权重图与原始特征图逐元素相乘，对特征图进行空间上的加权，以增强关键区域的特征表达能力。

3.2　改进损失函数

损失函数[19]用于衡量模型预测结果与真实标签之间的差异，从而指导模型的训练过程。在语义分割任务中，常用的损失函数包括交叉熵损失函数(Cross-Entropy Loss)、Dice损失函数等。采用混合损失函数，将交叉熵损失函数和Dice损失函数相结合。

交叉熵损失函数用于衡量模型预测的类别概率分布与真实标签之间的差异。对于语义分割任务，设p_{ij}为模型预测的第i个像素属于第j类的概率，y_{ij}为第i个像素的真实标签($y_{ij}\in\{0, 1\}$)，则交叉熵损失函数L_{CE}的计算公式为

$$L_{CE}=-\sum_{i=1}^{N}\sum_{j=1}^{K}y_{ij}\ln(p_{ij}) \tag{5}$$

式中：N为图像的像素总数，K为类别数。

交叉熵损失函数能够促使模型正确区分每个像素的类别，但对于类别不平衡的情况较为敏感。

Dice损失函数用于衡量预测分割结果与真实标签之间的重叠程度。Dice系数的计算公式为

$$D_{\text{Dice}} = \frac{2|P \cap G|}{|P| + |G|} \tag{6}$$

式中：P为预测的分割掩码，G为真实的分割掩码。

Dice损失函数L_{Dice}定义为

$$L_{\text{Dice}} = 1 - D_{\text{Dice}} \tag{7}$$

Dice损失函数在处理前景和背景的不平衡问题时表现出良好的鲁棒性，能够促使模型更准确地分割出目标区域。

通过将交叉熵损失函数和Dice损失函数进行加权求和，可以得到混合损失函数L，计算公式为

$$L = \alpha L_{\text{CE}} + (1 - \alpha) L_{\text{Dice}} \tag{8}$$

式中：α为权重系数。通过调整α的值，可平衡2种损失函数的作用。实验中经多次尝试，最终确定α=0.6。预测分割与真实分割效果对比如图9所示。

3.3 实验验证

实验采用与2.3节中同样的方法设置数据集，并对每张图像进行精细的人工标注，标注内容包括压力仪表的指针、刻度盘、数字显示等元素的分割掩码。

实验使用PyTorch深度学习框架实现改进算法。选用Xception作为DeepLabv3+模型的骨干网络。在训练过程中，采用Adam优化器，初始学习率设置为0.000 1，训练总轮数为100个epoch。在训练过程中，引入学习率衰减策略，每20个epoch将学习率降低为原来的0.5倍。

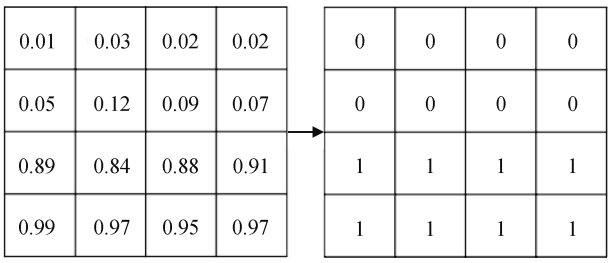


图9 预测分割与真实分割效果对比

Fig.9 Comparison between Predicted Segmentation and Ground Truth Segmentation Effects

采用像素准确率(Pixel Accuracy，PA)、平均交并比(Mean Intersection over Union，mIoU)、F1分数和参数量(Params)作为评价指标。

改进前后算法效果对比如表2所示。从实验结果可以看出：引入注意力机制后，mIoU提升2.7%，PA与F1分数分别增长1.2%和1.6%，但参数量增加15.6%。改进损失函数在保持模型复杂度的前提下，使mIoU提高了1.4%。同时引入注意力机制和改进损失函数后，PA提升了5.8%，mIoU提升了7.4%，F1分数提升了0.07。这表明模型在分割准确性和完整性方面有了明显的改善，但增加了模型复杂度；在处理复杂背景和光照变化的图像时，改进后的模型能够更准确地分割出压力仪表的各个元素，减少了误分割和漏分割的情况。

表2 元素分割算法效果对比

Tab.2 Effect comparison of element segmentation algorithms

模型	PA/%	mIoU/%	F1分数	Params/M
原始DeepLabv3+	83.5	71.2	0.76	4.5
引入注意力机制的DeepLabv3+	86.2	74.5	0.79	5.2
改进损失函数的DeepLabv3+	87.1	75.3	0.80	4.5
引入注意力机制和改进损失函数的DeepLab3+	89.3	78.6	0.83	5.2

4 系统实现

4.1 系统总体架构设计

基于OCR技术的压力仪表现场计量系统主要由硬件层、数据处理层、OCR识别层和应用层组成。

系统的硬件层利用高分辨力、宽动态范围和强抗干扰能力的仪器在不同的光照条件下清晰地采集压力仪表的图像，并对采集到的图像数据进行处理和分析，该层主要包括图像采集设备和计算设备。

数据处理层的主要任务是预处理采集到的原始图像，包括对图像进行增强、去噪、定位和裁剪等操作，以提高图像质量并实现特征的精准提取。具体手段包括直方图均衡化、自适应滤波、

高斯滤波、中值滤波以及裁剪原始图像等。

OCR识别层是系统的核心部分，负责对预处理后的图像进行字符识别。该层采用改进的OCR算法，结合深度学习技术和包括字符分割、特征提取和分类识别在内的传统图像处理方法，分离图像中的字符并识别提取其特征以完成分类，从而提高字符识别的准确率和鲁棒性。

应用层是系统的交互操作界面，需要同时面向用户和管理人员，同时提供友好易上手的用户界面，并支持数据查询、统计和分析等功能。此外，还包含便于管理人员进行系统配置、图像采集和结果查看等操作的数据管理模块，方便在实际生产活动中对提取到的识别结果进行存储和管理。

4.2　模块详细设计

如图10所示，硬件选型测试装置的图像采集模块能够控制工业级高清摄像机，依据用户设定的参数，定时采集压力仪表的图像数据，并将采集到的图像传输至数据处理层进行预处理。该装置的相关参数详见表3。

图10　硬件选型测试装置图

Fig.10　Diagram of hardware selection testing device

图像预处理模块对采集到的原始图像进行增强、去噪、定位、裁剪等一系列的处理。该模块可以采用直方图均衡化方法对图像的亮度和对比度进行调整，提高图像的清晰度以实现图像增强效果，或使用高斯滤波方法去除图像中的噪声干扰，平滑图像。此外，根据压力仪表的特征，使用改进后的Fast R-CNN方法，准确定位仪表区域并将其从原始图像中裁剪出来。

表3　硬件选型测试装置

Tab.3　Hardware selection testing device

硬件设备	关键参数
工业相机	镜头视野范围：67 mm × 100 mm；特征分辨力：15 μm × 15 μm；最小特征的像素点数：*FpX* 和 *FpY* 为3 pixels；像元尺寸：2.4 μm × 2.4 μm；帧速率达5 fps；分辨力达2 000万像素。
计算设备	英特尔酷睿Ultra 5处理器125 H
远心镜头	像圈直径：11.0 mm； 工作距离：181.8 mm； 像侧分辨力：5 μm。
LED环形光源	光衰预测：35 000 h亮度仍然维持原来的75%；使用寿命：约50 000 h。

OCR识别模块对预处理后的图像使用改进后的DeepLabv3+模型进行字符分割，随后通过CNN提取字符的特征信息，再将提取的特征信息输入到预训练的分类器中，完成字符的分类和识别，实现完整的字符识别过程。

用户交互界面模块采用直观易用的图形化界面设计，并支持传统且被大众广泛接受的鼠标和键盘操作。用户交互界面如图11所示。

图11　用户交互界面

Fig.11　User interaction interface

数据管理模块支持数据的查询、统计、分析等功能，采用数据库技术，将识别结果存储到数据库中以实现对识别结果的存储和管理。

4.3　OCR识别过程

1）图像预处理

采集到的图像可能由于工业现场的光照条件限制而出现亮度不均、对比度低等问题。为改善图像质量，可以采用直方图均衡化方法对图像进

行增强处理。直方图均衡化通过对图像的灰度直方图进行调整，使图像的灰度分布更加均匀。需要计算图像的灰度直方图和累积分布函数(Cumulative Distribution Function，CDF)，并根据CDF对图像的灰度值进行映射，从而得到增强后的图像。图像预处理效果如图12所示。

图12 图像预处理效果

Fig.12 Effect of image pre-processing

预处理完成后，对表盘进行检测定位，使用改进后的Fast R-CNN算法实现对仪表区域的准确定位并将其截取出来。目标检测效果如图13所示。

图13 目标检测效果

Fig.13 Object detection effect

2）字符分割

字符分割可将图像中的字符分离出来，以便后续的特征提取和识别。采用投影法对字符进行分割。投影法是一种基于字符在水平和垂直方向上的投影分布进行分割的方法，其具体实现步骤如下：①计算图像在水平方向上的投影分布；②根据投影分布，确定字符的边界；③根据字符的边界，将字符从图像中分割出来。改进模型对表盘的分割效果如图14所示。

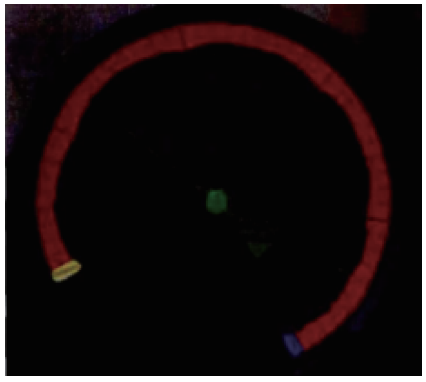


图14 改进模型对表盘分割效果

Fig.14 Segmentation effect of the improved model on the dial

3）特征提取

采用卷积神经网络实现字符特征信息的提取，以便后续的分类和识别。CNN是一种深度学习模型，具有自动提取图像特征的能力。为此，需要构建一个包括卷积层、池化层、全连接层等结构的CNN模型，并使用训练数据集对CNN模型进行训练，使其学习字符的特征信息。然后将分割后的字符图像输入到训练好的CNN模型中，以提取字符的特征信息。指针轮廓如图15所示。

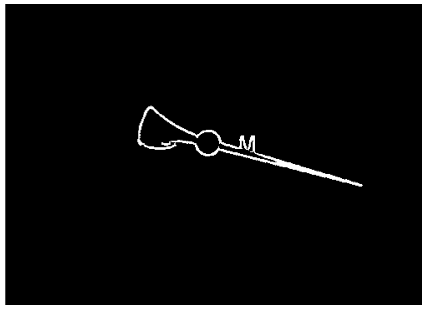


图15 指针轮廓图

Fig.15 Pointer contour diagram

4）分类识别

分类识别是OCR识别流程的最后一步，其目的是将提取的特征信息输入到预训练的分类器中，根据提取的特征信息对字符进行分类和识别。分类器可以采用支持向量机、神经网络等方法。在本系统中，采用神经网络作为分类器。具体而言，首先需要构建一个包括输入层、隐藏层和输出层等的神经网络模型，然后使用训练数据集对其进行训练，使其学习字符的分类信息。最后将提取的特征信息输入到训练好的神经网络模型中，完成对字符的分类和识别。

4.4 实验验证

本文构建了一个包含1 000张涵盖了不同的光照条件、仪表类型和安装姿态的压力仪表图像的数据集，并在每张图像上都人工标注了压力仪表上的数字读数，以此来验证系统的性能。

实验使用Python编程语言和TensorFlow深度学习框架实现系统功能，并按照7∶2∶1的比例将数据集划分为训练集、验证集和测试集。在训练过程中，采用随机梯度下降(Stochastic Gradient Descent，SGD)优化算法对模型训练100轮，并将学习率设置为0.001。

实验结果以准确率(识别正确的字符数占总字符数的比例)、召回率(识别正确的字符数占实际字符数的比例)、F1分数(准确率和召回率的调和平均数)作为评价指标。系统实验结果如表4所示。

表4　系统实验结果

Tab.4　System experimental results

单位：%

准确率	召回率	F1分数
94.5	95.8	96.1

实验结果表明：基于OCR技术的压力仪表现场计量系统在识别准确率、召回率和F1分数等评价指标上均达到了较高的水平，能够满足工业现场压力仪表计量的需求。

5　总结

本文围绕基于CNN的压力仪表OCR识别系统展开研究，主要分析了从仪表定位、字符分割到OCR识别等一系列问题。通过改进Fast R-CNN算法、引入数据增强、更换特征提取网络，实现了复杂环境下压力仪表的准确定位；通过在DeepLabv3+模型中引入通道注意力机制和空间注意力机制，并采用混合损失函数，提高了压力仪表元素分割算法的性能；最终成功实现了基于OCR技术的压力仪表现场计量系统。

在未来的研究中，算法优化仍是关键攻坚方向。需要致力于突破极低光照、强电磁干扰等极端工况下字符识别瓶颈，进一步提升微小字符和异形字符的识别准确度。此外，还需拓展OCR技术的应用边界，尝试适配更多类型和规格的工业仪表(如涵盖液位计、流量计、温度计等)，构建一个统一且通用的智能仪表计量平台。这将为计量自动化、智能化提供强大动力，创造更大的经济与社会效益。

参考文献

[1] 杨军，张力，李新良．动态计量技术发展中的几个关键问题[J]．计测技术，2021，41(2)：8-12.
YANG J，ZHANG L，LI X L. Several primary problems in the development of dynamic metrology[J]. Metrology & Measurement Technology，2021，41(2)：8-12.（in Chinese）

[2] 周曼，刘志勇，鲁乾鹏，等．基于OCR的数字仪表自动识别在工业现场中的应用[J]．仪器仪表用户，2021，28(1). DOI:10. 3969/j. issn. 1671-1041. 2021. 01. 006.
ZHOU M，LIU Z Y，LU Q P，et al. Application of automatic recognition of digital instruments based on OCR in industrial field[J]. Instrument and Meter Users，2021，28(1). DOI：10. 3969/j. issn. 1671-1041. 2021. 01. 006.（in Chinese）

[3] 何欣．计量强基康斯特仪表公司智能制造在前进[J]．中国计量，2023(5)：43-44.
HE X. Strengthening metrology foundation：const instrument′s intelligent manufacturing advances[J]. China Metrology，2023(5)：43-44.（in Chinese）

[4] 黄良武，王文涛．视觉检测技术在发动机装配线上的应用[J]．制造业自动化，2024，46(3)：217-223.
HUANG L W，WANG W T. Application of visual inspection technology in the engine assembly line[J]. Manufacturing Automation，2024，46(3)：217-223.（in Chinese）

[5] 谌鎏．商汤科技公司发展战略研究[D]．武汉：华中科技大学，2024.
CHEN L. Research on the development strategy of sensetime group[D]. Wuhan：Huazhong University of Science and Technology，2024.（in Chinese）

[6] 梅永，庄建军．基于Canny边缘检测的图像预处理优化算法[J]．信息技术，2022(1)：75-79.
MEI Y，ZHUANG J J. Optimization algorithm for image preprocessing based on Canny edge detection[J]. Information Technology，2022(1)：75-79.（in Chinese）

[7] 郭庆梅，刘宁波，王中训，等．基于深度学习的目标检测算法综述[J]．探测与控制学报，2023，45(6)：10-20，26.
GUO Q M，LIU N B，WANG Z X，et al. Review of target detection algorithms based on deep learning[J]. Journal of Detection and Control，2023，45(6)：10-20，26.（in Chinese）

[8] 杨晨，侯志强，李新月，等．基于CNN-Transformer双模态特征融合的目标检测算法[J]．光子学报，2024，53(3)：280-293.

YANG C, HOU Z Q, LI X Y, et al. Target detection algorithm based on CNN-Transformer dual-modal feature fusion[J]. Acta Photonica Sinica, 2024, 53(3): 280-293. (in Chinese)

[9] 张敬勋, 张俊虎, 赵宇波, 等. 基于字符分割和LeNet-5网络的字符验证码识别[J]. 计算机测量与控制, 2023, 31(7): 271-277.
ZHANG J X, ZHANG J H, ZHAO Y B, et al. Character CAPTCHA recognition based on character segmentation and LeNet-5 network[J]. Computer Measurement & Control, 2023, 31(7): 271-277. (in Chinese)

[10] 张彩珍, 李颖, 康斌龙, 等. 基于深度学习的模糊车牌字符识别算法[J]. 激光与光电子学进展, 2021, 58(16): 259-266.
ZHANG C Z, LI Y, KANG B L, et al. Fuzzy license plate character recognition algorithm based on deep learning [J]. Progress in Laser & Optoelectronics, 2021, 58(16): 259-266. (in Chinese)

[11] 马玲, 罗晓曙, 蒋品群. 基于模板匹配和支持向量机的点阵字符识别研究[J]. 计算机工程与应用, 2020, 56(4): 134-139.
MA L, LUO X S, JIANG P Q. Research on dot matrix character recognition based on template matching and support vector machine [J]. Computer Engineering and Applications, 2020, 56(4): 134-139. (in Chinese)

[12] 齐兴斌, 赵丽, 耿海军, 等. 基于改进Faster R-CNN的域自适应红外目标检测方法[J]. 计算机工程与设计, 2024, 45(10): 2994-3001.
QI X B, ZHAO L, GENG H H, et al. Domain adaptive infrared target detection method based on improved Faster R-CNN[J]. Computer Engineering and Design, 2024, 45(10): 2994-3001. (in Chinese)

[13] 张光钱, 周广利, 黄飞, 等. 面向3D目标检测的多模态生成式图像数据增强的研究[J]. 重庆理工大学学报(自然科学), 2024, 38(10): 13-20.
ZHANG G Q, ZHOU G L, HUANG F, et al. Research on multimodal generative image data augmentation for 3D object detection[J]. Journal of Chongqing University of Technology (Natural Science), 2024, 38(10): 13-20. (in Chinese)

[14] 方宇伦, 陈雪纯, 杜世昌, 等. 基于轻量化深度学习VGG16网络模型的表面缺陷检测方法[J]. 机械设计与研究, 2023, 39(2): 143-147.
FANG Y L, CHEN X C, DU S C, et al. Surface defect detection method based on lightweight deep learning VGG16 network model[J]. Mechanical Design and Research, 2023, 39(2): 143-147. (in Chinese)

[15] 曲英伟, 刘锐. 基于YOLOv5-MobileNetV3算法的目标检测[J]. 计算机系统应用, 2024, 33(7): 213-221.
QU Y W, LIU R. Object detection based on YOLOv5-MobileNetV3 algorithm[J]. Computer System Applications, 2024, 33(7): 213-221. (in Chinese)

[16] 慕涛阳, 赵伟, 胡晓宇, 等. 基于改进的DeepLabV3+模型结合无人机遥感的水稻倒伏识别方法[J]. 中国农业大学学报, 2022, 27(2): 143-154.
MU T Y, ZHAO W, HU X Y, et al. Method for identifying rice lodging based on improved DeepLabV3+ model combined with UAV remote sensing[J]. Journal of China Agricultural University, 2022, 27(2): 143-154. (in Chinese)

[17] 孙义博, 张文靖, 王蓉, 等. 基于通道注意力机制的行人重识别方法[J]. 北京航空航天大学学报, 2022, 48(5): 881-889.
SUN Y B, ZHANG W J, WANG R, et al. Pedestrian re-identification method based on channel attention mechanism[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(5): 881-889. (in Chinese)

[18] 张宸嘉, 朱磊, 俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20): 64-72.
ZHANG C J, ZHU L, YU L. Review of attention mechanism in convolutional neural network[J]. Computer Engineering and Applications, 2021, 57(20): 64-72. (in Chinese)

[19] 陈科圻, 朱志亮, 邓小明, 等. 多尺度目标检测的深度学习研究综述[J]. 软件学报, 2021, 32(4): 1201-1227.
CHEN K Q, ZHU Z L, DENG X M, et al. Review of deep learning for multi-scale object detection[J]. Journal of Software, 2021, 32(4): 1201-1227. (in Chinese)

(本文编辑: 李成成)

第一作者: 王晶星(1997—), 女, 助理工程师, 主要研究嵌入式测量仪器。

通信作者: 王丽(1980—), 女, 高级工程师, 主要研究压力测量仪器。