

基于改进U2-Net的指针式仪表读数方法

王 奇, 李 莉, 王 震, 王树云

天津职业技术师范大学电子工程学院, 天津

收稿日期: 2024年8月28日; 录用日期: 2024年9月21日; 发布日期: 2024年9月29日

摘要

当前工业指针式仪表读数过程中存在特殊环境下依赖人工和推理精度低等问题, 本文提出一种基于改进U2-Net的指针式仪表读数方法。针对目前仪表识别算法推理精度差和模型参数数量过多的问题, 将U2-Net编码阶段的RSU4和RSU5的最深层的两个卷积更换成深度可分离卷积, 并在每个RSU的编码阶段后加入了ECA注意力模块, 使模型更好地关注指针和刻度区域, 提高指针和刻度的识别精度。本文在收集到的数据集上进行评估, 通过对比实验表明, 相较于SegNet、Deeplabv3+及U2-Net方法, 本文改进的模型查准率达到94.58%, 针对两种量程25 MPa和1.6 MPa的压力仪表读数引用误差达到1.012%, 具有较好的性能表现。

关键词

指针式仪表, U2-Net, 自动读数, 仪表识别

Reading Method of Pointer Instrument Based on Improved U2-Net

Qi Wang, Li Li, Zhen Wang, Shuyun Wang

School of Electronic Engineering, Tianjin Vocational and Technical Normal University, Tianjin

Received: Aug. 28th, 2024; accepted: Sep. 21st, 2024; published: Sep. 29th, 2024

Abstract

At present, there are some problems in the reading process of industrial pointer instruments, such as relying on manual work and low reasoning accuracy. This paper proposes a reading method for pointer instruments based on improved U2-Net. Aiming at the problems of poor reasoning accuracy and too many model parameters in the current instrument identification algorithm, the deepest two convolutions of RSU4 and RSU5 in the U2-Net coding stage are replaced by deep separable convolutions, and the ECA attention module is added after each RSU coding stage, which made the model

pay more attention to the pointer and scale area and improved the recognition accuracy of pointer and scale. In this paper, the collected data sets are evaluated. Compared with SegNet, Deeplabv3+ and U2-Net methods, the accuracy of the improved model in this paper reaches 94.58%, and the reference error of pressure instruments with two measuring ranges of 25 MPa and 1.6 MPa reaches 1.012%, which has good performance.

Keywords

Pointer Instrument, U2-Net, Automatic Reading, Instrument Identification

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

复杂的工业环境中存在辐射、毒性、高温等恶劣环境，需要借助仪器对生产条件进行检查以确保安全。传统上，获取的数据通常是由人工读取的，此类方法能够从各种形状、形式和风格的复杂仪表中获得精确的读数。然而，人工方法耗时且费力。相比之下，依靠巡检机器人和计算机视觉技术进行自动抄表具有重要的现实意义[1]。目前工业用指针式仪表具有造价较低、适用于多种现实环境、便于维修更换等特点，在某些工业生产中具有广泛且不可替代的应用。然而，复杂的工业环境增加了人工读表的困难，目前语义分割和读数算法已达到较低识别误差，可直接用于指针式仪表读数，因此研究一种基于人工智能技术的指针式仪表自动读数方法迫在眉睫。

目前，指针式仪表的识别方法主要分为传统方法和深度学习方法。传统的指针式仪表识别方法主要依赖于传统的图像处理技术和特定的算法。房桦等人提出了一种用于设备仪表检测机器人的仪表识别算法[2]，使用尺度不变特征变换方法来检测仪表区域，利用快速霍夫变换检测两条指针直线去除噪音干扰，定位指针精确位置和指向角度，完成指针读数。Gao JW 等人利用随机样本一致性拟合仪表圆弧，通过加速稳健特征(Speeded up Robust Features, SURF)方法检测仪表区域，利用霍夫变换方法寻找指针，最终通过角度法计算仪表读数[3]。为了更好地满足实际情况下变电站的检测需求，Kucheruk V 等人在 ORB 特征检测的基础上做了算法的改进，通过匹配特征点快速确定指针中心，并最终通过霍夫变换实现仪表指针读数的读取[4]。然而，传统的识别方法在面对复杂的自然环境时，可能会受到背景环境改变的影响，导致识别精度下降。基于深度学习的指针式仪表识别方法可以通过训练模型学习仪表的刻度分布和指针的特征，实现表盘和指针的自动定位，并根据学习到的特征计算指针的角度和读数。万吉林等人[5]在 FasterR-CNN 的目标检测和 U-Net 图像分割的基础上，对指针式仪表的读数进行了自动识别。金爱萍等人[6]结合了 YOLOv5 与 U-Net 两种深度学习模型的优点，旨在解决不同背景和距离条件下的指针识别难题。U2-Net [7]作为一种先进的语义分割模型，在医学图像分割等领域取得了显著成效。然而，直接将其应用于指针式仪表读数识别时，仍存在推理精度待提升等问题。

针对以上问题，本文提出一种基于改进 U2-Net 的指针式仪表读数方法。本文设计了基于加入深度可分离卷积配合 ECA 注意力模块的特征融合模块，自适应分割和学习指针仪表的刻度和指针等特征。为了解决模型编码过程中模型参数过多的问题，本文将部分模块的较深层次部分卷积替换为深度可分离卷积，同时在编码阶段后加入 ECA 注意力模块用于关注小目标区域。通过将 U2-Net 进行改进，将 U2-Net 的优点进行有效结合，从而提高指针式仪表检测识别网络的推理精度和速度。

2. 相关技术和理论

在本文研究的指针仪表提取特征过程中，更加重视细节信息和全局对比性信息。鉴于 U2-Net 网络是针对显著性目标检测任务提出的，能够将图像中最具有吸引力的目标提取出来[8]，故本文采用 U2-Net 语义分割模型对指针式仪表进行识别。

U2-Net [7]的每个 RSU 块都是类似于 U-Net [9]的网络架构，结构如图 1 所示。

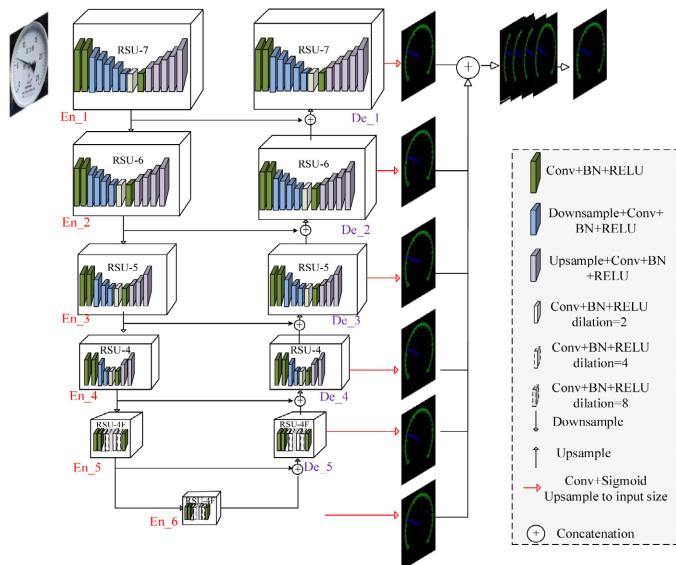


Figure 1. Network structure diagram of U2-Net

图 1. U2-Net 网络结构图

网络的主体类似于 U-Net 的网络结构，在大的 U-Net 中，每一个小的 block 都是一个小型的类似于 U-Net 的结构。整体结构由一个编码器和一个解码器组成，编码器负责提取图像的特征，而解码器则负责将特征映射回原始图像尺寸在编码器阶段，每通过一个模块后都会下采样 2 倍，在解码器阶段，每通过一个 block 前会上采样 2 倍。

U2-Net 的 En_1 和 De_1 阶段，使用 RSU-7 模块，RSU-7 的网络结构图如图 2 所示。En_2 和 De_2 阶段使用 RSU-6，相对于 RSU-7 少了一个下采样卷积以及上采样卷积的部分，RSU-6 block 只会下采样 16 倍，而 RSU-7 block 下采样了 32 倍。

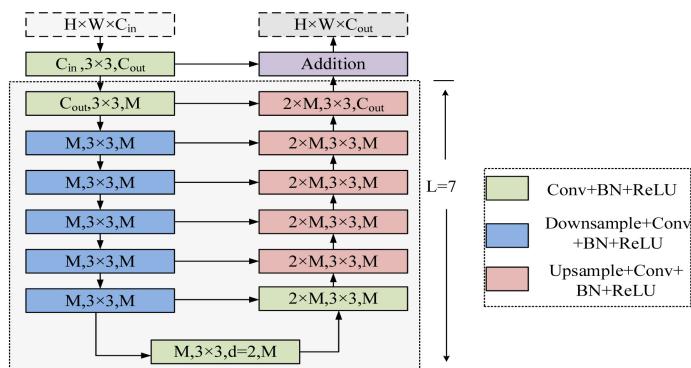


Figure 2. Network structure diagram of RSU-7

图 2. RSU-7 网络结构图

图像经过 En_1-En_4 下采样处理后对应特征图的高与宽相对较小，如果再进行下采样操作不免会丢失有用性数据，为了保留上下文信息，En_5、En_6、De_5 不再进行下采样，且 RSU-4F 的模块中，将下采样、上采样结构换成了膨胀卷积。

3. 仪表检测识别方法

本文提出的基于改进 U2-Net 的指针式仪表语义分割算法，能有效提高模型的查准度，如图 3 所示，模型包含 2 个模块，分别为：仪表语义分割模块、读数后处理。

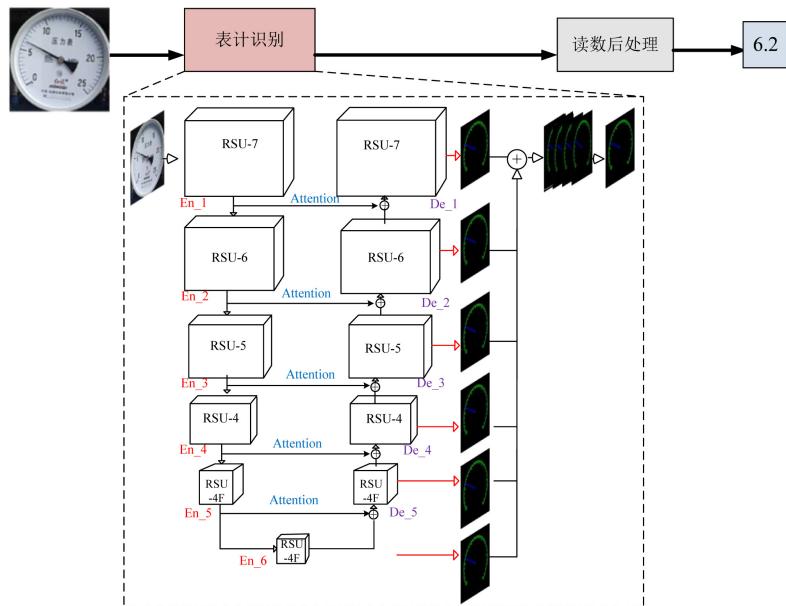


Figure 3. Structure diagram of this method
图 3. 本文方法结构图

3.1. 基于改进 U2-Net 的表计识别

U2-Net 基于 Encode-Decide 编码形式，编码过程中存在细节特征损失和模型参数过多问题，针对该问题，将编码阶段的 RSU4 和 RSU5 模块的最深层次的两个普通卷积更换为深度可分离卷积[10]，降低参数数量并提升对细节特征的学习能力。

深度可分离卷积(Depthwise Separable Convolution, DSConv)由逐深度卷积和逐点卷积组成，深度卷积用于提取空间特征，逐点卷积用于提取通道特征。深度可分离卷积在特征维度上分组卷积，对每个通道进行独立的逐深度卷积，并在输出前使用一个 1×1 卷积将所有通道进行聚合。

深度可分离卷积和普通卷积的参数量和计算量对比如下：

$$\text{参数量比值: } \frac{\text{深度可分离卷积}}{\text{标准卷积}} = \frac{D_K \times D_K \times M + M \times N}{D_K \times D_K \times M \times N} = \frac{1}{N} + \frac{1}{D_K^2}$$

$$\text{计算量比值: } \frac{\text{深度可分离卷积}}{\text{标准卷积}} = \frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F} = \frac{1}{N} + \frac{1}{D_K^2}$$

其中， D_K 表示卷积核大小，M 和 N 分别表示逐深度卷积和逐点卷积的卷积核大小，一般 N 较大，故 $\frac{1}{N}$ 可忽略不计。

深度可分离卷积的结构如图 4 所示。

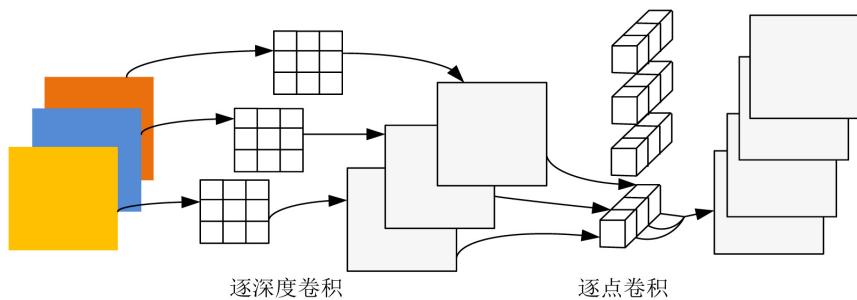


Figure 4. Depth-separable convolution structure diagram
图 4. 深度可分离卷积结构图

从以上公式和结构图中，可以观察到，按如上改进后的总参数量比普通卷积的总参数量更低。本次改进通过提高计算效率和优化网络结构，减少了过拟合的风险和计算资源的消耗。

为了帮助模型在特征融合时更好地关注感兴趣区域，提高指针分割精度，将各编码 RSU 拼接下采样时加入 ECA 注意力模块，以 RSU-7 为例，改进模型如图 5 所示。

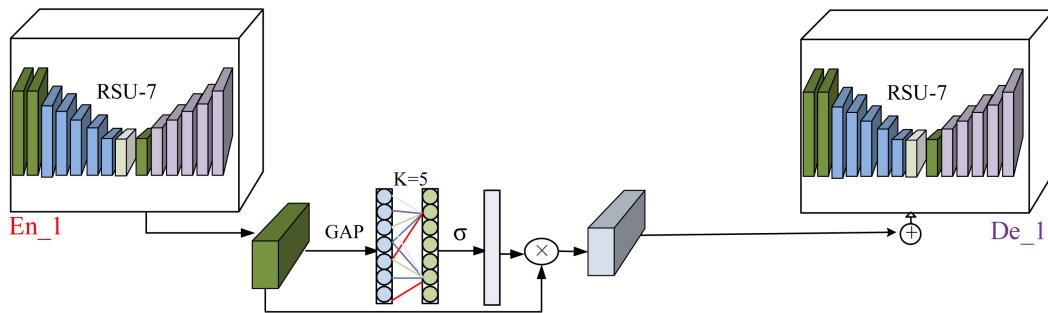


Figure 5. Feature fusion combined with ECA attention module
图 5. 结合 ECA 注意力模块的特征融合

ECANet [11] 的核心思想是在卷积操作中引入通道注意力机制，以捕捉不同通道之间的关系，从而提升特征表示的能力。通道注意力机制的目标是自适应地调整通道特征的权重，使得网络可以更好地关注重要的特征，抑制不重要的特征。在每个解码阶段 RSU 模块后加入 ECA 注意力模块，可以确保模型在不增加过多参数和计算成本的情况下，有效地增强网络的表征能力。

3.2. 读数后处理

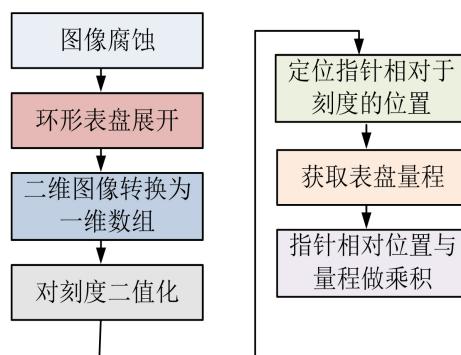


Figure 6. Flow chart of reading post-processing
图 6. 读数后处理的流程图

使用改进后的 U2-Net 模型将训练集仪表的指针和刻度分割出来后，经过读数后处理得到最终的读数。读数后处理的流程图如图 6 所示。

主要包含以下步骤：

第一步，对语义分割的预测类别图进行图像腐蚀操作，以达到刻度细分的目的。

第二步，把环形的表盘展开为矩形图像，根据图像中类别信息生成一维的刻度数组和一维的指针数组。

第三步，接着计算刻度数组的均值，用均值对刻度数组进行二值化操作。

第四步，根据刻度分割得到总体刻度格数，大于 40 格归类为 25 MPa 类型(50 格)，小于 40 格归类为 1.6 MPa (32 格)，根据指针分割得到指针相对位置，最后将指针相对位置与分类量程做乘积得到表盘的读数。

本文在对指针仪表上标注了刻度和指针两个类别，并将指针和刻度分开输入到语义分割网络中，故将图像划分成双通道分别处理，取出掩码图第一通道的特征图，像素等于 1 的为指针，像素等于 2 的为刻度。输入图像经过模型语义分割和图像处理后的掩码图，如图 7 所示。

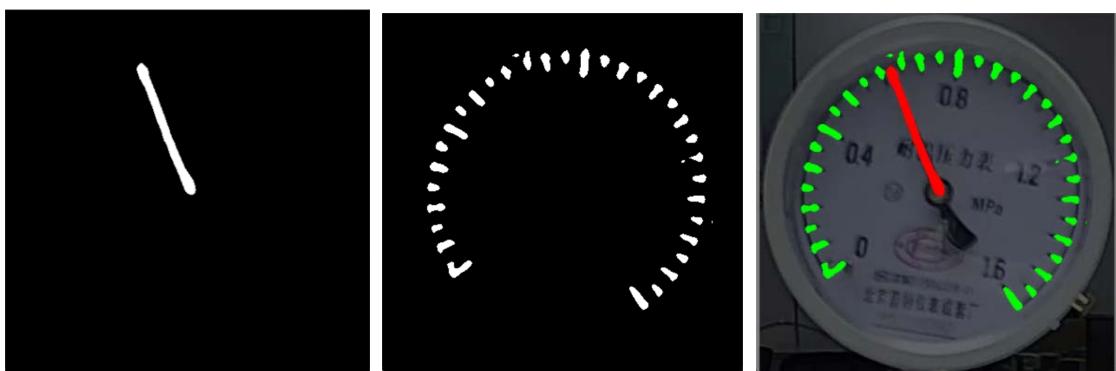


Figure 7. Mask diagram of pointer and scale

图 7. 指针和刻度的掩码图

通过对比输入图像与模型推理后的图像，可以明显看出模型的推理效果相当出色。模型对于表盘内的指针和刻度等关键特征的提取非常精确，充分证明了经过改进后的 U2-Net 模型在仪表识别任务中展现出了卓越的处理能力，验证了模型的有效性。

3.3. 损失函数

交叉熵损失是一个常用于多分类问题的损失函数。交叉熵损失函数使用在单个类中，使得多分类任务计算损失变得更加高效和简单，如公式 1 所示。

$$H(p, q) = -\sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (1)$$

x_i 是变量， $p(x)$ 、 $q(x)$ 为 x 的两个概率分布，交叉熵越小， p 和 q 两分布越相似。本文设置背景类别的权重是 1，指针和刻度类别的权重为 1.5。

4. 实验结果

4.1. 参数设置和数据集

本文使用的操作系统为 Windows11，GPU 选用 NVIDIA GTX-1080，深度学习框架为 Pytorch 2.2.2。表计识别网络使用工厂实拍数据图像，本次实验共包含了 2 种量程的仪表，分别为 1.6 MPa 和 25 MPa，

训练集为 374 张，测试集为 37 张。部分数据集图像及掩码图可视化如图 8 所示。



Figure 8. Data set and mask map
图 8. 数据集及掩码图

4.2. 评估指标

本文采用查准率(Precision, P)、平均交并比(Mean Intersection over Union, mIoU)、召回率(Recall, R)作为评价指标，对图像分割模型结果进行分析，公式如下：

$$P = \frac{p_{ii}}{p_{ii} + p_{ji}} \quad (2)$$

$$\text{mIOU} = \frac{1}{n+1} \sum_{i=0}^n \frac{p_{ii}}{\sum_{j=0}^n p_{ij} + \sum_{j=0}^n p_{ji} - p_{ii}} \quad (3)$$

$$R = \frac{p_{ii}}{p_{ii} + p_{ij}} \quad (4)$$

其中， n 表示类别总数， p_{ii} 指像素类别为 i 被预测为类别 i 的总数量， p_{ij} 指像素类别为 i 被预测为类别 j 的总数量， p_{ji} 指像素类别为 j 被预测为类别 i 的总数量。

为判定仪表读数的准确性，选择引用误差率来判断模型的推理效果，引用误差越小，模型推理越准确。引用误差 e 的表达式如下：

$$e = \frac{|\text{value} - \text{value}'|}{l} \quad (5)$$

其中，模型推理读数为 value ，仪表的人眼读数为 value' ，仪表的分类量程为 l 。

4.3. 消融实验

为了验证所提方法的有效性，本文对测试集进行了消融实验，并使用查准率 P 、平均交并比 mIOU、召回率 R 来评价各模型，实验结果如表 1 所示。

- 1) U2-Net: U2-Net 进行指针仪表数据集训练推理。
- 2) U2-Net + DSConv: 利用 DSConv 改进的 U2-Net 进行指针仪表数据集训练推理。
- 3) U2-Net + ECA 注意力: 利用 ECA 注意力改进的 U2-Net 进行指针仪表数据集训练推理。
- 4) 本文方法: 使用改进的 U2-Net 进行指针仪表数据集训练推理。

Table 1. Ablation experimental results**表 1.** 消融实验结果

序号	方法	P	mIOU	R
1	U2-Net	90.13%	82.83%	89.18%
2	U2-Net + DSConv	92.21%	83.70%	90.55%
3	U2-Net + ECA 注意力	93.62%	85.37%	90.26%
4	本文方法	94.73%	85.49%	92.65%

从表 2 可知, 在 U2-Net 模型中引入 ECA 注意力特征融合方法并将部分普通卷积替换为 DSConv 后, 本文方法相较于原始的 U2-Net 模型和模型单个改进方法, 三项评价指标均有明显提升。相较于原始模型, 本文方法的查准率 P 提高了 4.6%, 平均交并比 mIOU 提升了 2.66%, 召回率 R 增加了 3.47%。说明模型在识别指针和刻度时, 更加准确地捕捉到了目标区域, 减少了误判和误识别的情况, 为后续读数误差的降低打下了基础。本文改进不仅提高了模型的分割效果, 还为实际应用中的指针式仪表识别带来了更加准确和稳定的解决方案。

4.4. 对比实验

为了证明本文所提方法的真实有效性, 将本文方法与 SegNet [12]、Deeplabv3+ [13]、U2-Net [7] 等经典的语义分割模型进行了比较, 具体结果见表 2。

Table 2. Contrast experimental results**表 2.** 对比实验结果

序号	方法	p	mIOU	R
1	SegNet	88.23%	80.97%	88.60%
2	Deeplabv3+	89.13%	81.89%	88.85%
3	U2-Net	90.13%	82.83%	89.18%
4	本文方法	94.73%	85.49%	92.65%

根据表 3 的对比数据可知, 相较于目前常用的语义分割网络 SegNet、Deeplabv3+ 和 U2-Net 相比, 本文方法评价指标明显提升。在与 SegNet 的查准率数据比较中, 两者差距为 6.5%, 比 DeepLabv3+ 高出 5.6 个百分点, 本文方法同样展现出优越的性能。与原始模型 U2-Net 相比, 本文方法在阈值为 0.01 的条件下, 精度比 U2-Net 提高了 4.6 个百分点。综上所述, 本文提出的模型方法在指针式仪表识别的查准率、平均交并比、召回率方面均展现出显著的优势。

4.5. 模型训练效果

本实验共包含 2 种量程的压力仪表, 分别为 25 MPa 和 1.6 MPa, 测试集共 37 张图片。为准确评估仪表读数的精确性, 本实验采用引用误差率作为评价模型推理效果的重要指标。引用误差率通过衡量模

型预测值与真实值之间差距来说明模型推理的准确性，值越小，模型的推理效果越准确，预测值与真实值之间的偏差越小。本文详细记录了部分推理结果，呈现在表 3 中，同时计算了用不同语义分割模型的训练权重计算平均引用误差 e ，结果见表 4。

Table 3. Model reasoning result**表 3.** 模型推理结果

序号	人眼读数	推理读数	引用误差 e (%)
1	1.11	1.1827	4.544
2	6.25	6.3368	0.347
3	6.4	6.3559	0.176
4	0.63	0.6113	1.117
5	0.613	0.6235	0.656
6	6.2	6.0919	0.432

Table 4. Inference accuracy and average citation error of different semantic segmentation models (%)**表 4.** 不同语义分割模型的推理精度和平均引用误差(%)

序号	方法	P	e (%)
1	SegNet	88.23%	3.207
2	Deeplabv3+	89.13%	3.431
3	U2-Net	90.13%	1.686
4	本文方法	94.73%	1.012

由表 3、表 4 可知，本文提出的模型展现出了较低的引用误差率，表明本文所提模型在识别指针式仪表时，能够更准确地提取出关键信息，并给出接近真实值的预测读数结果。再一次证明了本文提出的模型在仪表识别任务中具有出色的推理效果，能够准确地识别出仪表的读数，为相关领域的应用提供了可靠的技术支持。

5. 结论

针对指针式仪表在工业生产中遇到的复杂环境下人工读数难以实现和读数误差较大的问题，本文提出了一种基于人工智能的指针式仪表自动读数方法，通过改进的 U2-Net 语义分割模型分割指针和刻度，为了在提高识别精度的同时减少计算量，改进普通卷积并加入注意力机制。经验证，该模型能够高效地将各仪表的指针和刻度分割出来，并且能够通过分割效果识别仪表类型和指针相对位置从而计算读数，最终改进的 U2-Net 模型的查准率达到 94.58%，针对两种量程 25 MPa 和 1.6 MPa 的压力仪表读数引用误差达到 1.012%，相较于原 U2-Net 模型和目前主流语义分割模型均有显著提升。

然而，尽管本文方法在比较清晰的数据集情况下表现出色，但其读数后处理阶段仍采用传统的角度法，并且目前不能适用于更广泛的其他类型指针式仪表，在处理反光和异物遮挡等复杂图像时，可能导致处理效果不佳。为了进一步提高识别性能，我们将对仪表读数的计算方法进行深入研究，并积极探索和开发新的读数方法，以应对各种复杂的实际应用场景。

参考文献

- [1] 杨兆昭. 基于深度学习的指针式仪表识别方法研究[D]: [硕士学位论文]. 西安: 西安石油大学, 2023.
- [2] 房桦, 蒋涛, 李红玉, 等. 一种适用于智能变电站巡检机器人的双针仪表读数的识别算法[J]. 山东电力技术,

- 2013(3): 9-13+69.
- [3] Gao, J., Xie, H., Zuo, L. and Zhang, C. (2017) A Robust Pointer Meter Reading Recognition Method for Substation Inspection Robot. 2017 *International Conference on Robotics and Automation Sciences*, Hongkong, 26-29 August 2017, 43-47. <https://doi.org/10.1109/icras.2017.8071914>
- [4] Kucheruk, V., Kurytnik, I., Kulakov, P., et al. (2018) Definition of Dynamic Characteristics of Pointer Measuring Devices on the Basis of Automatic Indications Determination. *Archives of Control Sciences*, **28**, 401-418.
- [5] 万吉林, 王慧芳, 管敏渊, 等. 基于 Faster R-CNN 和 U-Net 的变电站指针式仪表读数自动识别方法[J]. 电网技术, 2020, 44(8): 3097-3105.
- [6] 金爱萍, 袁亮, 周德勤, 等. 基于 YOLOv5 和 U-Net 的指针式仪表读数识别方法[J]. 仪表技术与传感器, 2022(11): 29-33.
- [7] Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O.R. and Jagersand, M. (2020) U2-Net: Going Deeper with Nested U-Structure for Salient Object Detection. *Pattern Recognition*, **106**, Article 107404. <https://doi.org/10.1016/j.patcog.2020.107404>
- [8] 舒小锋, 吴晓红, 卿粼波, 等. 基于改进 U-2Net 的岩石薄片图像分割[J]. 计算机系统应用, 2024, 33(2): 159-165.
- [9] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M. and Frangi, A.F., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Springer, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [10] Howard, A.G., Zhu, M., Chen, B., et al. (2017) Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications. <https://arxiv.org/abs/1704.04861>
- [11] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W. and Hu, Q. (2020). ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 11531-11539. <https://doi.org/10.1109/cvpr42600.2020.01155>
- [12] Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 2481-2495. <https://doi.org/10.1109/tpami.2016.2644615>
- [13] Chen, L., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H. (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 833-851. https://doi.org/10.1007/978-3-030-01234-2_49