中国传媒大学本科毕业论文（设计）中期报告表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 毕设题目 | 基于计算机视觉的仪表识别技术研究和开发 | | | | |
| 学生姓名 | 史骏诚 | 学号 | 2020211123012 | 专业 | 人工智能 |
| 指导老师 | 苗方 | | | | |
| 中期报告内容：   1. 引言   在现代工业生产和电力系统管理中，模拟仪表扮演着至关重要的角色，它们被广泛用于监控和记录各种物理量，如电流、压力和温度。然而，在许多情况下，这些仪表处于高危、难以接近或人员密集的环境中，如变电站和化工厂。在这些环境中，人工读取仪表不仅耗时、劳动强度大，而且可能存在安全风险。因此，开发一种自动化读取模拟仪表的方法变得尤为重要，它不仅可以提高数据收集的效率和准确性，还能确保人员安全。  尽管计算机视觉技术在许多领域取得了显著进步，但在自动仪表读取方面仍面临一系列挑战。其中包括仪表面板的多样性、环境光照条件的变化、以及仪表指针和刻度的识别难题。此外，现有的自动仪表读取系统往往无法处理低质量图像，如模糊、反射或仪表破损的情况，这限制了它们在实际应用中的可靠性和广泛性。  鉴于此，本研究的主要目标是通过现有的计算机视觉技术，开发一个既高效又准确的自动仪表读取系统。具体而言，本研究致力于解决仪表检测、图像对齐和数字识别等关键技术问题，旨在实现对各种类型和质量仪表图像的准确解析。通过这项研究，我们希望提高自动化仪表读取技术的适应性和鲁棒性，以满足复杂工业环境中的实际需求。   1. 相关工作   在过去的几十年里，仪表读取一直是工业自动化领域关注的焦点。传统的仪表读取方法主要依赖于人工操作，工作人员需要定期检查并记录仪表的读数。这种方法不仅耗时耗力，而且容易受到人为错误的影响，特别是在恶劣或危险的环境中，如高辐射或高温条件下。此外，由于需要连续监控，人工读取无法实现实时数据更新，限制了其在现代自动化和智能监控系统中的应用。  随着技术的发展，自动仪表读取技术开始逐步兴起。早期的自动化尝试主要集中在基于规则的图像处理方法，例如边缘检测和模式匹配。虽然这些方法在特定条件下能够工作，但它们通常缺乏灵活性和鲁棒性，特别是在面对仪表款式和环境条件多样化时。近年来，随着深度学习技术的兴起，计算机视觉在仪表识别方面取得了显著进步。然而，即使是最先进的方法也仍然面临着处理复杂场景、多样化的仪表类型和不同光照条件的挑战。  在此背景下，YOLO-v8的应用提供了新的解决方案。作为最新一代的物体检测技术，YOLO-v8在速度和准确性方面都有显著提升。它能够在接近实时的速度下处理图像，同时保持高精度，这使其非常适合于需要快速响应的仪表读取任务。更重要的是，YOLO-v5支持多尺度训练和推理，这意味着它能够更好地适应不同大小和形状的仪表。此外，YOLO-v8的强大性能和灵活性使其能够更好地应对复杂背景和变化的光照条件，这是传统自动仪表读取技术难以解决的问题。因此，将YOLO-v8应用于仪表检测不仅可以提高检测的准确率，还可以提高系统的整体效率和可靠性。   1. 实验设计   数据集介绍：采用了专门设计用于评估自动仪表读取系统性能的MC1296数据集。该数据集包含1296张高质量的仪表图像，这些图像涵盖了不同类型的仪表面板，包括但不限于电压表、压力表和温度计。数据集中的图像来源于多样化的环境背景，模拟了实际工业现场的复杂情况，包括不同的光照条件、仪表面板的磨损或脏污，以及多种视角和距离。这种多样性确保了测试的全面性和挑战性，使我们能够准确评估模型在现实世界应用中的表现。   |  |  |  |  | | --- | --- | --- | --- | | Dataset\_task | Train\_size | Test\_size | Annotations | | M\_detection | 1036 | 260 | mb | | M\_alignment | 1028 | 247 | co | | M\_reading | 739 | 185 | psn |   表1：“mb”、“co” 和 “psn”代表仪表边界框、坐标偏移量以及指针/刻度/数字掩码和数字  实验原理及设置：  对于指针仪表而言，主流的识别原理主要通过提取指针位置，零刻度位置，量程数字。最后通过三者之间的角度换算。    假设绿色箭头和红色箭头角度为，蓝色箭头（指针位置）和红色箭头(零刻度)角度为，量程是(本图为0.04) 。那么最后识别的结果就是。  基于此原理设计了三阶段的仪表识别系统，分别为仪表检测，仪表矫正，仪表识别，这三个阶段由三个主要模块构成，分别是空间转换模块（STM）[7]、值获取模块（VAM）[8]和基于YOLO的检测模块（YDM）。首先，YDM负责检测图像中的表计并进行裁剪，以便于后续处理。接着，STM模块通过学习的方式自动校正图像，以获得表计的正面视图，从而克服了传统透视变换方法在处理点匹配时的缺点。最后，VAM模块整合了表计组件检索和表数值识别的过程，通过深度学习模型直接从校正后的图像中识别和计算表计的读数。这一整套流程旨在模拟人类的识别行为，提高了读表的准确性和效率。   1. 方法论   1.仪表检测：在这一阶段，采用了基于最新一代YOLO-V8的改进检测模型（YDM）。  1.1.预处理和增强：实现包括了对输入图像的预处理和数据增强步骤。预处理包括调整图像大小、归一化等。数据增强技术，如随机裁剪、旋转、缩放和颜色扭曲，用来增加模型的泛化能力并减少过拟合  1.2.多尺度训练：YOLO-v8型进行多尺度训练，这意味着它可以检测不同 大小的仪表。这一点对于仪表检测特别重要，因为仪表的大小和视角在不同的图像中可能有很大差异。  1.3.训练过程：模型在准备好的数据集上进行训练，使用标准的损失函数（如交叉熵损失）和优化器（如Adam或SGD）。训练过程包括调整模型权重以最小化预测和真实标注之间的差异。    2.仪表对齐：引入了改进的空间变换网络（STM）[7]来解决仪表图像对齐的问题。  2.1方法概述：STM采用了一种隐式学习方法，直接从数据中学习图像间的空间变换关系，而非传统STN中的显式参数回归方法。这种设计避免了复杂的点匹配步骤，减少了模型的预处理需求。  2.2结构和实现：STM结构基于改进的STN，使用ResNet18作为特征提取的主干网络。它通过训练学习确定从原图到对齐图之间的空间映射。与传统STN相比，STM不直接输出变换矩阵，而是输出一组变换指示，例如图像中的关键位置偏移，这些指示随后被用于计算最终的同质性变换。    传统的STN空间变换网络（中）和改进的空间变换网络（下）  3.仪表识别：在识别阶段，采用了值获取模块（VAM）[8]，这是一个综合的框架，用于检索仪表上的关键组件并识别数字。  3.1.深度学习网络：VAM使用ResNet18作为其主干网络来提取仪表图像的特 征，并利用特征金字塔网络（FPN）进行特征的多尺度融合，确保了对不同大小和形状仪表组件的有效识别。  3.2.组件检索分支：此分支利用语义分割方法来精准定位仪表的关键组件，如指针和刻度。通过两个独立的1×1卷积层生成的指针图和关键刻度图来实现，同时使用Dice损失函数来优化分割性能。  3.3.数字识别分支：VAM通过一个专门的分支来识别仪表上的数字，该分支能够处理文本识别任务，并利用在线难例挖掘（OHEM）策略来提高关键数字检测的准确率。关键数字识别是通过将检测到的数字区域转换成固定大小并通过双向LSTM进行序列识别来实现的。     1. 结果分析  |  |  |  |  | | --- | --- | --- | --- | | Model | AP50(%) | AP75(%) | FPS | | Liu.et al[1] | 91.3 | 89.5 | 4.3 | | YOLO[2] | 90.0 | 88.2 | 6.7 | | YOLO-v8 | 98.6 | 97.1 | 12.4 |   表2：不同仪表检测方法的定量结果，AP为平均精度。  实验结果显示，基于YOLO-v8的检测模型（YDM）在仪表检测任务上具有较高的准确率和速度，显著超过了传统的物体检测方法。   |  |  |  |  | | --- | --- | --- | --- | | Method | Rel(%) | Ref(%) | FPS | | None | 5.91 | 1.20 | - | | Perspective transform[1] | 1.72 | 0.23 | 10 | | STN[3] | 3.4 | 0.95 | 44 | | STM[7] | 1.7 | 0.26 | 50 |   表3：不同仪表对准方法的定量结果，“rel”是 平均相对误差，“ref”是平均参考误差。  此外，改进的空间变换网络（STM）在仪表图像对齐方面表现出色，与传统的对齐方法相比，它可以更快速地处理更广泛的角度偏差，从而为数字识别提供了更准确的基础。   |  |  |  |  | | --- | --- | --- | --- | | Method | Avenue | Ref(%) | Ref(%) | | Zheng et al.[4] | Measurement(2016) | 10.32 | 0.91 | | Gao et al.[5] | ICRAS(2017) | 9.34 | 0.67 | | He et al.[6] | ICIST(2019) | 1.85 | 0.30 | | Liu et al.[1] | Measurement(2020) | 1.77 | 0.24 | | VAM[8] | - | 1.70 | 0.26 |   表4：不同抄表识别方法的定量结果”Rel“是平均相对误差，”Ref“是平均参考误差。  值获取模块（VAM）在数字识别任务上同样表现出色，准确识别了绝大多数仪表读数，即使在面对复杂背景和模糊图像时也能保持较高的准确度。  整体而言，该实验方法在各项指标上表现优异，证明了其在自动化仪表读取方面的有效性和优越性。这些结果不仅展示了单个模块的性能，还验证了整个系统的协同效应，提高了自动仪表读取技术的实际应用潜力。   1. 参考文献 2. Liu, Y., Liu, J., Ke, Y.: A detection and recognition system of pointer meters in substations based on computer vision. Measurement 152, 107333 (2020) 3. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 779–788 (2016) 4. aderberg, M., Simonyan, K., Zisserman, A., et al.: Spatial transformer networks. Advances in neural information processing systems 28 (2015) 5. Zheng, C., Wang, S., Zhang, Y., Zhang, P., Zhao, Y.: A robust and automatic recognition system of analog instruments in power system by using computer vision. Measurement 92,413–420 (2016) 6. Gao, J.W., Xie, H.T., Zuo, L., Zhang, C.H.: A robust pointer meter reading recognition method for substation inspection robot. In: 2017 International Conference on Robotics and Automation Sciences (ICRAS). pp. 43–47. IEEE (2017) 7. He, P., Zuo, L., Zhang, C., Zhang, Z.: A value recognition algorithm for pointer meter based on improved mask-rcnn. In: 2019 9th International Conference on Information Science and Technology (ICIST). pp. 108–113. IEEE (2019) 8. Yang, C., Xie, W. and Zisserman, A. (2022) It’s about time: Analog clock reading in the wild, arXiv.org. Available at: https://arxiv.org/abs/2111.09162 9. Shu, Y. et al. (2023) Read pointer meters in complex environments based on a human-like alignment and recognition algorithm, arXiv.org. Available at: https://arxiv.org/abs/2302.14323   日期： 学生（签字）： | | | | | |
| 开题结果 |  | | | | |
| 指导老师（签字）：  年 月 日 | | | | | |