TECH OFFER

Video-Level Assisted Data Labelling For Industrial Applications



KEY INFORMATION

TECHNOLOGY CATEGORY:
Infocomm - Artificial Intelligence
Infocomm - Video/Image Analysis & Computer Vision

TECHNOLOGY READINESS LEVEL (TRL): TRL6
COUNTRY: SINGAPORE
ID NUMBER: TO174664

OVERVIEW

Existing publicly available datasets, such as COCO, are built from the ground up to be general-purpose and therefore lack domain specificity. When such public datasets are used to train deep learning models for industrial use-cases and applications, e.g. detection of electronic components, they often result in sub-par performance caused by the disparity between objects typically found in industrial environments and data residing in public datasets. This disparity requires significant effort in pixel-level supervision (annotation), where each pixel, per frame, has to be annotated manually to make up for the difference in training data to improve model performance

This solution is a deep-learning-based technique for instance segmentation in industrial environments intended to reduce the effort cost of annotation from pixel-level to video-level. With instance segmentation, the goal is not just to detect and localise objects within a scene, but also to determine the different classes and number of instances (or recognising more of the same type objects as different). This aids scene understanding and the resulting model can be deployed for productivity measurement or process improvement. Incremental learning is used to ensure that only the parts of the model that need to be updated with new

© COPYRIGHT 2023 - IPI. ALL RIGHTS RESERVED.

data are changed, thus reducing the amount of time taken for re-training and model updates.

TECHNOLOGY FEATURES & SPECIFICATIONS

Data collection

- The data regarding a target object (object to be classified) is collected via depth cameras, one at a time
- For static objects, the camera is rotated around the target object, whereas for mobile objects, the camera is fixed statically, and multiple viewpoints are used to capture the moving object from a variety of angles
- Multiple clean background images (without any objects) are also captured for accurate segmentation

Pseudo labels

Instead of annotating every frame within the video, pseudo-pixel-level labels for each video frame are generated through 4 steps:

- Image-based weakly supervised segmentation
- 3D registration-based weakly supervised segmentation
- Optical flow-based mask generation
- Merging of each segmented layer and refinement

Labels derived from the video-level are then applied to the combined segments as pseudo-labels.

Real-time inference with incremental learning

Leveraging the existing classification capability of a neural network that has been pre-trained on a COCO dataset to classify 80 original COCO classes, incremental learning is used to build a new classifier that can classify a new target object e.g. cargo container, circuit board, plastic bottle etc. The output of the original classifier and generated pseudo labels from the previous step are combined and used to train this new classifier. This new classifier is generated separately in order to avoid affecting the original model's generic classification capability.

POTENTIAL APPLICATIONS

This solution is applicable for various industrial applications such as factories, warehouses and cargo terminals. Additionally, it can be deployed as part of any automated system that requires computer vision based instance segmentation/object recognition or on robots and existing surveillance cameras.

UNIQUE VALUE PROPOSITION

In comparison with existing methods which are often developed on general-purpose public datasets and require pixel-level annotation for new training data to be added, this solution abstracts data annotation to the video-level, while producing similar performance in instance segmentation results. Additionally, the costs of development and implementation are greatly reduced since the bottleneck of annotation is minimised.

© COPYRIGHT 2023 - IPI. ALL RIGHTS RESERVED.