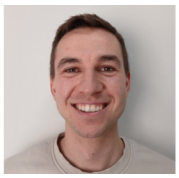
Crate Detection

Students

Urs Lacher

Lukas Leuzinger



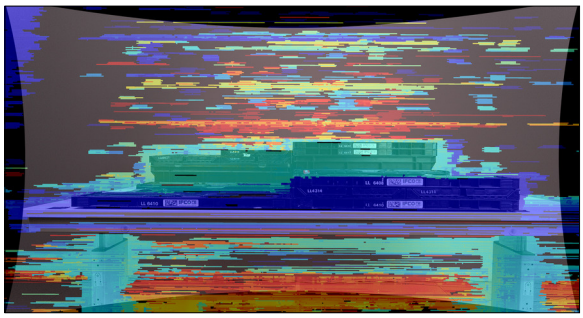
Roman Weiss

Problem: In a logistics warehouse, crates bearing goods are being sent and returned empty and folded frequently. Recollected crates are returned on a pallet containing mixed crate types. They can be dirty, and their stacking can show displacements. Classification and counting must currently be done manually. With our thesis we wanted to explore the possibility of automating the process of detecting and counting these crates.

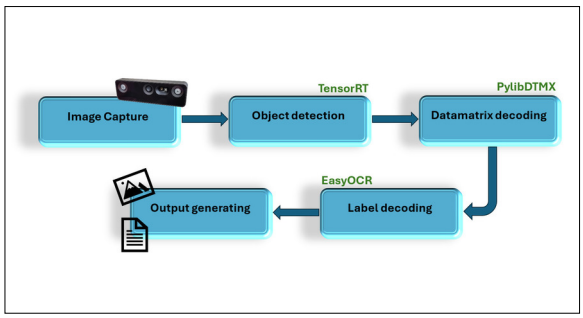
Approach / Technology: To detect the crates, computer vision techniques were used. The performance was optimized and tested on a set of 10 crate instances with four different crate-types. With a "YOLOv11l Oriented Bounding Boxes" Model from Ultralytics the crates, labels and data matrices on the crates should be detected. The model was trained on a dataset containing 250 images. To improve the performance on our use case with the specified object classes, we used augmentations to virtually expand the test set. Further optimization of the model was made through hyperparameter tuning. The data matrices were used to uniquely identify the crates and their type. If the data matrix can not be detected or decoded, the label on each crate is read through “EasyOCR” to identify the type of crate. Multiple crate-stacks appear on one image. To distinguish them, it is necessary to get depth information. The chosen approach was to use an OAK D Pro W camera. This device has an RGB camera with 12MP resolution and 2 mono 1MP cameras to determine the depth information using triangulation. The image and depth information had to be mapped together to only count crates on the stack nearby to the camera. A Jetson AGX Orin was chosen for the machine learning tasks, because of the suited hardware and the software support from NVIDIA.

Result: A pipeline was setup to proceed the RGB and depth information of the camera through object detection and our algorithms to output the detected crates with all their information in a JSON-file. On the NVIDIA Jetson AGX orin and the given set of crates, the output could be generated within 1.6 seconds. On the test set of 10 images with 78 crates, there were 85.9% of the crates identified with the correct type. For the 11 wrong detected crates, there were six times an error with OCR decoding and five times due to missing YOLO detections.

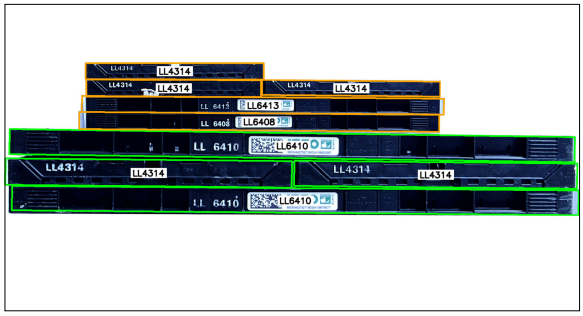
The crates in the front and the back and the colored overlay as the depth measurement (distance: camera to crates)
Own presentment



This flow chart shows graphically how the internal process works.
Own presentment



Shown are the recognised boxes. The valid ones in green (distance = 70cm), the invalid ones in orange (distance > 70cm)
Own presentment



Advisors

Prof. Dr. Martin Weisenhorn, Marc Benz

Subject Area

Image Processing and Computer Vision