Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 105 (2022) 1–6

www.elsevier.com/locate/procedia

29th CIRP Life Cycle Engineering Conference

Vision-Based Screw Head Detection for Automated Disassembly for Remanufacturing

Simon Mangold*, Christian Steiner, Marco Friedmann, Jürgen Fleischer

Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

* Corresponding author. Tel.: +49-1523-950-2634. *E-mail address:* simon.mangold@kit.edu

Abstract

Remanufacturing is commonly perceived as a promising field for future challenges such as resource efficient production. For an economic operation of remanufacturing facilities, an automation of the currently manual labor is mandatory. Thus, the automation plays a vital role in order to realize high rates of re-utilization and therefore a significant reduction of waste. Screw connections allow for non-destructive dismantling and are commonly used connection elements. Especially the automation of the disassembly step is a key element as products from the field are of unknown specification upon feeding to the remanufacturing line due to alterations during their life cycles. State of the art solutions for automated disassembly lack flexibility to adapt to different products and product conditions. This contribution presents a highly flexible approach for the localization and classification of screws in electric motors. The presented system utilizes a tool equipped industrial robot with an integrated eye-in-hand vision system and an industrial computer. The system is able to locate and classify six different types of screw heads of varying sizes using machine learning approaches in order to adapt the robot's end-effector. Because of the presented hardware concept the system depends upon a minimum of constraints concerning the presentation of objects. This paper compares different network architectures and peripheral settings and presents the most suitable solution to the use case. A dataset consisting of six classes of different screw heads was created to train neural networks to detect screws in an experimental set-up consisting of metal blocks holding different screws of diverse types and conditions. Results are validated on two different electric motors from the automotive sector being processed on an automated disassembly line.

Keywords: Automated disassembly; Remanufacturing; Screw Head Detection

1. Introduction

In order to keep planet earth vigorous and for the maintenance of our societies, humankind needs to achieve sustainability and climate neutrality. Therefore private lifestyles and industrial operations need to change. One measure on this behalf in production technology is the so-called remanufacturing, where used and partly worn-out parts are being reworked and brought to "as new" conditions [1]. Within this process chain, the dismantling of products is mandatory. Nowadays, especially disassembly for remanufacturing is being conducted mainly in manual labor [2]. However, for future large-scale applications with high numbers of remanufactured parts, it is mandatory to automate this task. At the same time, the automation of the disassembly for remanufacturing implies many challenges for production systems. In comparison to manufacturing production lines with similar parts and connectors with small alterations and variations, disassembly lines address parts and connectors with large variations. In the case of remanufacturing, the parts have passed at least one prior life cycle where many inaccessible alterations have changed the conditions of the parts and connectors. An automated dismantling system thus needs to adapt to the individual conditions of the processed parts. Among others, this is considered as one of the main challenges for production systems [3]. The disassembly system needs to adapt on the physical layer as in special the selection and usage of tools [4]. Within this contribution, an image-processing system for the detection of screw heads is presented. The system is applied to the adaption of robot tools. To maintain the industrial applicability, the system is embedded in an industrially relevant use case. The addressed cores are two different supplementary motors from the automotive discipline.

10.1016/j.procir.2022.02.001

2. Related Work

Since bolted connections are commonly used connectors, vision based screw detection has found interest from research in prior works. In a contribution by Vongbunyoung et al., the automatic detection of connective components is described as a highly desirable task for a disassembly system [5]. Bolted connections are described as rather easy to detect in comparison to other commonly used connective components such as rivets, tape or glue. This is as screw heads are highly standardized and generally visible components. Nonetheless, there are two major challenges highlighted and shown by a case study of Liquid-crystal display (LCD) panels:

- Small size of connectors poses high demands to the utilized camera system.
- High variety because of different materials, corrosion or lighting conditions pose high demands on the flexibility of a visual detection system.

The presented screw detection system focuses on the distinction between product components and connector elements but does not include the classification task to different screw heads. Another related image processing system is proposed in [6]. The authors derive a visual intelligence scheme for the disassembly of hard drives. The underlying detection algorithm was first introduced in [7]. In the later publication, the screw detection is accomplished in a two step approach: First, candidates were generated under the assumption that all screw heads have a circular shape. Therefore an OpenCV Hough Transformation is utilized to detect circles. Found candidates are then further examined since not all found circular shapes are screw heads. The contribution examines several classifiers to the task and derives a model which achieves an accuracy of 99%. However, both described systems are only capable of distinguishing between the two classes "screws" and "artifacts". There is no further subdivision into different kinds of screw heads.

In the work of [8] the authors describe an architecture to distinguish between four different types of screw heads in different sizes. The final classification achieves an accuracy of 97%. The pipeline's overall accuracy is limited by the candidate generation done by searching for circular structures in the image using the OpenCV Hough Transformation. It was shown that this step succeeds in approximately 75 % of all attempts. This resulted in an average precision (AP) of 0.757 for the complete pipeline. The authors suggest the exploration of alternative methods for the detection of region of interests (ROI). It is also noted that most misclassifications can be directly linked to how Hough cuts regions.

A possible alternative approach to detect crosshead screws is described by [9]. The proposed system relies on a digital single-lens reflex camera (DSLR) camera connected to a computer that performs the screw detection using several image processing techniques and the Faster-RCNN with Inception Version 2 deep learning architecture. It was shown that the image processing steps had an impact on the precision and recall of the deep learning model. The highest precision and recall rates of the system amount to 91.8 % and 83.6%. This was achieved by utilizing a combination of pre-processing, object detection and visual reasoning. The authors suggest the extension of the described system to other screw types.

Another notable mention goes to [3]. The developed system is able to perform all actions to completely disassemble electromechanical devices such as hard disk drives. Screw detection in this system also relies on the OpenCV Hough Transformation in order to detect possible locations of screws. Afterwards a binary classification algorithm is used to distinguish screws from or any other circular object. This classifier achieved an evaluation accuracy of 93,6 %. The evaluation of the disassembly process concluded that 10 out of 180 actions failed with 9 of them being unscrew actions.

Recent developments in the wide field of machine learning create a variety of possibilities for applications in the industrial environment. As summarized in [10] there are two major categories of object detectors. *One-Stage* or *Two-Stage Object Detectors* are distinguished by the way they generate the class probabilities and location coordinates of an object. In general one-stage object detectors are considered to have a lower detection accuracy than two-stage object detectors. The main advantage of a one-stage object detection is the high detection speed.

One of the representatives of one-stage object detectors is YOLO (You Only Look Once). YOLOv3 is the latest version developed by the original developer Joseph Redmon [11]. YOLOv4 was then released in 2020 and improved YOLOv3's AP and frames per second (FPS) by 10% and 12% respectively [12].

YOLOv5 is the newest version of the YOLO architecture[1] developed by Glenn Jocher[2] using Python's PyTorch framework.

Summarizing state of the art solutions, one can state that the required variety in the detection of screw heads for the application in automated disassembly for remanufacturing has not been solved. The existing algorithms and approaches are not containing the desired amount of different screw heads for the application. On the other hand, machine learning based approaches inherently provide mandatory flexibility. The YOLO architecture in the newest version provides the required framework and is utilized in the following.

3. Methodology

3.1. Overview

The scope of the presented system is the adaption of a robot tool towards a presented object containing screw heads. Therefore a camera system is aligned to capture the device. To address the autonomous adaption of the robot tools, a two step approach is used. At first, the screw head classification algorithm

[1] There has not yet been a publication on YOLOv5. However sourcecode and documentation can be accessed by: https://github.com/ultralytics/yolov5

[2] https://github.com/glenn-jocher

is applied. Based on this result, robot movements are executed to configure an appropriate tool for the disassembly task in the second step. This contribution focuses on the development and optimization of the screw detection part which is addressed by the following steps. Generated screw poses can subsequently be used to reach the screw head and conduct the dismantling steps.

- Localization of screws within the image
- Classification of screws into discrete classes

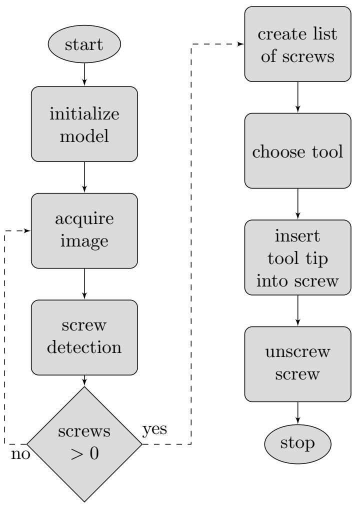

Fig. 1: Screw detection and disassembly workflow.

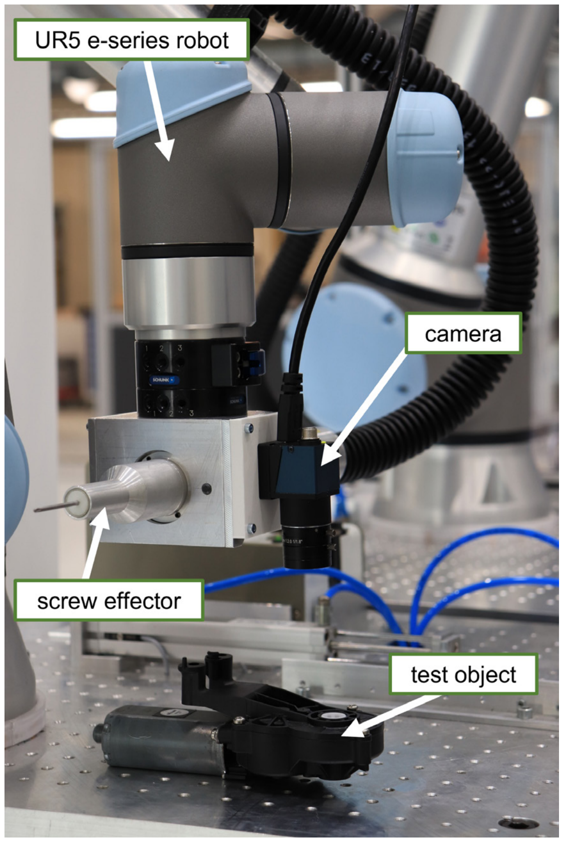

Fig. 2: Hardware components implemented as an eye-in-hand effector system during on-site validation.

3.2. Hardware

The focus of the contribution is on industrial applicability. Therefore, industrially relevant components are utilized.
The system consists of a Universal Robots UR5 e-series robot with an eye-in-hand camera system. Fig. 2 shows the combination of the camera and screw driver module attached to the robot flange.

A monochromatic camera with 1280x1024 pixels[3] equipped with a 6 mm lens[4] is used.
The object detection is deployed to a Siemens SIMATIC IPC227E[5]. Since computational power on this device is limited, it is necessary to achieve an acceptable runtime in order to allow for real world applications in a throughput driven environment.

3.3. Screw Detector

The localization and classification of screws in the perceived images is primarily implemented using the YOLOv5[6] object detection architecture. For the training, a dataset containing 550 images is created. Three different kinds of objects are included in the dataset to create a versatile system:

- Metal bars with equidistant bores holding screws of varying sizes, as seen in Figure 3(a) and (b).
- Electric window lifter motors with different conditions of connectors and parts (Fig. 3(c) and (d)).
- Starter motors for automotive combustion engines (Fig. 3(e) and (f)).

Within the data set the following six different classes of screw heads are included. All classes include the two sizes M3 and M6 respectively:

- Pozidriv
- Hex socket
- Slot drive
- Phillips
- Hex
- Torx

For the training dataset, no specified position and orientation of objects in the image were used. During the image acquisition for the training dataset, the camera is moved around the object. This leads to a large number of points of view and therefore in a desirable variety of images of a limited number of objects. Furthermore, different backgrounds, lighting conditions and apertures are used.
Both YOLOv5 screw detectors were trained for 500 epochs

[3] MER-134-93U3M USB3 camera by Daheng Imaging: https://www.get-cameras.com/USB-Camera-OnSemi-Python1300-MER-134-93U3M
[4] LCM-5MP-06MM-F2.0-1.8-ND1: https://www.get-cameras.com/LENS-C-mount-5MP-6MM-F2.0-1/1.8-LCM-5MP-06MM-F2.0-1.8-ND1
[5] SIMATIC IPC227E; Intel Celeron N2930; 8 GB RAM; 240 GB SSD; Industrial Edge
[6] Repository: https://github.com/ultralytics/yolov5
Official Documentation: https://docs.ultralytics.com/

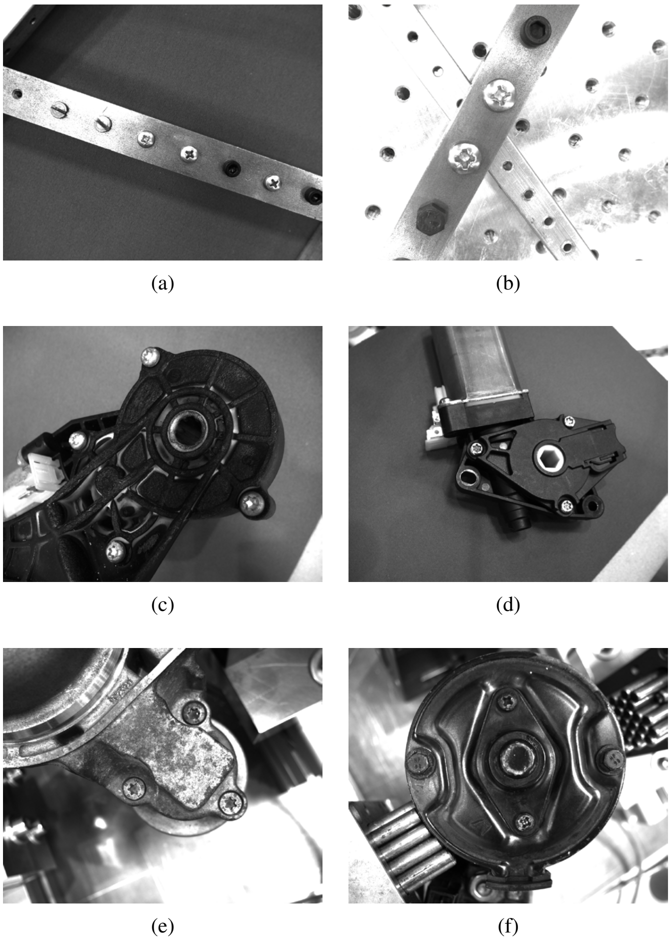

Fig. 3: Exemplary images contained in the dataset.

with a batch size of 32. The size of incoming images was reduced from the original size to 640 × 640 pixels. Table 1 shows the number of layers and number of parameters, such as the number of floating-point operations (GFLOPs). In order to allow reproducibility and an efficient training process, transfer learning using the `yolov5s.pt` and `yolov5m.pt` weights provided by ultralytics is conducted. The developed software package is integrated into the Siemens Industrial EDGE environment.

Detector	Layers	Parameters	GFLOPs
YOLOv5s	224	7,067,395	16.4
YOLOv5m	308	21,057,843	50.4

Table 1: Comparison of YOLOv5 detector architectures.

4. Experimental Results

4.1. Overall Precision

In order to evaluate the different detectors with respect to the present use case, different metrics were examined. The performance of the different detectors was performed using the Intersection over Union (IoU) method. Table 2 shows the achieved precision of different screw detectors.

Detector	mAP@0.5	mAP@0.5:0.95
YOLOv5s	0.984	0.834
YOLOv5m	0.980	0.826

Table 2: Experimental evaluation of different screw detectors.

It can be seen that there is no significant difference in overall precision. Even though the YOLOv5s version is a smaller and faster model it outperforms the YOLOv5m version on the used dataset. To evaluate the performance of the screw detector the confusion matrices (see Figures 4 and 5) are generated in order to get a further understanding on possible sources of errors.

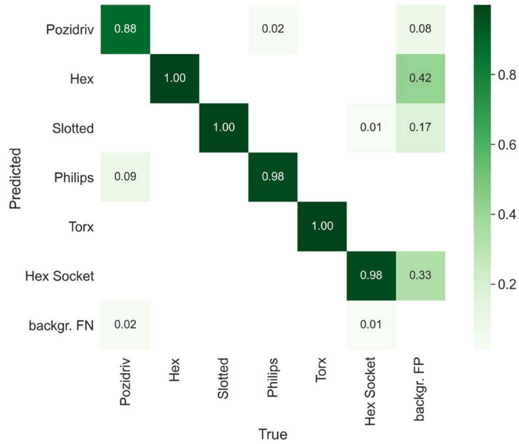

Fig. 4: Confusion matrix acquired for YOLOv5s on the screw detection dataset.

Fig. 5: Confusion matrix acquired for YOLOv5m on the screw detection dataset.

4.2. Runtime Evaluation

In addition to the precision metric, the runtime is regarded in special. It is defined as the average time of the pre-loaded detectors to acquire an image and achieve a result. Table 3 shows the results of the experimental runtime evaluation. Experiments were performed with 100 random images on the aforementioned Siemens SIMATIC IPC227E.

Detector	av. runtime [s]	rel. av. runtime
YOLOv5s	7.53	1.0
YOLOv5m	18.09	2.4

Table 3: Experimental runtime evaluation of two YOLOv5 screw detectors on Siemens SIMATIC IPC227E.

4.3. On-Site Validation

The entire software package containing the YOLOv5s model is deployed to the site of usage. The validation is conducted using unseen objects for the system, similar to the ones contained in the training dataset. A small motor (window lifter motor as in Figure 3 (c) and (d)) is repeatedly presented to the system with an alteration of the background, an alteration of the screw head type and a adaption of camera parameters. Likewise, a large motor (Starter motors as in Figure 3 (e) and (f)) is presented. The contained screw heads were partly damaged, dirty and or corroded. Table 4 contains the results of the test series.

Test Object	Present screws	Correct detections
20x Small motor	45	39
7x Large motor	17	15

Table 4: Experimental on-site validation on unseen objects

During the application of the system to the small motor, some errors occur. If the Torx head screws is changed to a Slot drive, the system falsely detects a Torx head. One may conclude that the system has learned alternative factors than the screw head form. Additionally there are False Positive results occurring. Some screw heads are being detected doubled. Despite these misdetections, the system appears rather robust against an alteration of the background and camera settings such as the focal length and the aperture setting. At the application of the large motor, there are fewer misdetections occurring. Yet there are False Positive errors occurring if the motor is only party visible in the image. All in all, the object detector is proven for robust functionality and has an acceptable overall runtime. However, it has to be mentioned that several influences on process quality could be observed.

Since it is not possible to automatically re-focus the camera optics, the positioning of the camera in relation to the object and therefore the image quality influence the outcome of the screw detection. The detection accuracy can be increased by using the robot to adjust the distance between camera and object in order to acquire a clear image. Screw detection on the electric motors works sufficiently reliable.

Another influence on the detection quality is the background of the images respectively the board under the cores. The detector is easily confused by the holes in the base plate (Figure 2) especially if the image has a poor quality. There are False Positive detections of hexagon socket screws observed.

5. Result Assessment

The confusion matrix and the classification results prove the general functionality of the approach and the screw detectors. Especially the distinction between the *Philips* and *Pozidriv* works precisely even though the screw heads appear rather similar to the human eye. Hexagon headed screws and hexagon socket screws are major sources for misclassifications. However, the biggest source of errors are False Positive detections. During the on-site validation it showed that these False Positive errors usually show a lower value of confidence provided by the screw detector. In a production environment this could be handled by applying a confidence threshold and ignoring detections that show a low confidence value.

As denoted in section 4.1, there is no significant advantage in terms of precision between the two applied classifiers.

Regarding the runtime evaluation, it can be seen that the larger screw classifier takes around twice as much time to evaluate a single image. In order to allow a productive usage of the developed package we advise the use of the YOLOv5s model at this stage of the development as it has significant performance advantages while maintaining a comparable accuracy as the YOLOv5m model.

The system is applied to scenes without any fixtures. In real environments, fixtures for the clamping of parts are used. Those may contain screws which could potentially lead to misclassifications. In this work, this effect is not considered. One approach could be the usage of image pre-processing steps such as filters for the repeating sections of the input images. In real production environments it is not guaranteed that screw heads face the camera plane as assumed in the contribution. Even though images within the presented training data set are taken from different angles, multiple views may be required for the detection of a screw in real production environments. Within this contribution only the ideal case is regarded and not-aligned perspectives are neglected. In real world applications one may conduct a classification into component levels and in a second step apply the presented screw detection in this contribution. The objects in the training dataset are screws of mainly good physical conditions. Yet in case of disassembly applications, the production system needs to work on damaged or worn out screws. It is not regarded profoundly in this contribution how the detector works on those objects. Further research should include this.

6. Conclusion

We present a system for the classification of screw heads into discrete classes corresponding to their physical shapes using state of the art object detectors. Together with the classification results, the location in the camera frame is derived. The system improves state of the art solutions as it is capable of classifying into multiple different classes. With the utilized hardware, we prove the application in an industrial environment. Especially the Industrial EDGE Environment of company Siemens can be denoted here. In future research, the correlation of the image quality to the screw detection results need to be assessed. Increased computational power would allow for more complex and presumably more accurate detectors. A larger dataset would augment the detection results and allow a better generalization and therefore a wider variety of possible applications of the system. This should be focused in further examinations. Experiments on the complete dismantling pipeline are not conducted since the contribution focuses on the image processing task solely. This task remains to future work. Furthermore, the influence of the training data on the outcome of the process, in special the reduction in variability will likely lead to an increase in overall precision.

Acknowledgements

The project AgiProbot is funded by the Carl Zeiss Foundation.

References

[1] Jawahir, I. S., Bradley, R., 2016. Technological Elements of Circular Economy and the Principles of 6R-Based Closed-loop Material Flow in Sustainable Manufacturing. Procedia CIRP 40, 103–108.

[2] Javier Ramírez, F., Aledo, J., Gamez, J., Pham, D., 2020. Economic modelling of robotic disassembly in end-of-life product recovery for remanufacturing. Computers & Industrial Engineering 142.1

[3] Klas, C., Hundhausen, F., Gao, J., Dreher, C., Reither, S., Zhou, Y., Asfour, T., 2021. "The KIT Gripper: A Multi-Functional Gripper for Disassembly Tasks" IEEE International Conference on Robotics and Automation (ICRA) 2021. Xi'an, China

[4] Fleischer, J., Gerlitz, E., Rieß, S., 2021. Concepts and Requirements for Flexible Disassembly Systems for Drive Train Components of Electric Vehicles. Procedia CIRP 98, 577-582.

[5] Vongbunyong, S., Chen, W. H., 2015. Disassembly Automation. Automated Systems with Cognitive Abilities. Springer International Publishing

[6] Yildiz, E., Brinker, T., Renaudo, E., Hollenstein, J., Haller-Seeber, S., Piater, J., Wörgötter, F., 2020. A Visual Intelligence Scheme for Hard Drive Disassembly in Automated Recycling Routines. Proceedings of the International Conference on Robotics, Computer Vision and Intelligent Systems (ROBOVIS 2020)

[7] Yildiz, E., Wörgötter, F., 2019. "DCNN-Based Screw Detection for Automated Disassembly Processes" 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS) 2019. Sorrento, Italy

[8] Yildiz, E., Wörgötter, F., 2020. DCNN-based Screw Classification in Automated Disassembly Processes. Proceedings of the International Conference on Robotics, Computer Vision and Intelligent Systems (ROBOVIS 2020), 61–68.

[9] Foo, G., Kara, S., Pagnucco, M., 2021. Screw detection for disassembly of electronic waste using reasoning and re-training of a deep learning model. Procedia CIRP 98, 666–671.

[10] Xiao, Y., Tian, Z., Yu, J., Zhang, Y., Liu, S., Du, S., Lan, X., 2020. A review of object detection based on deep learning. Multimedia Tools and Applications 79.11, 23729–23791

[11] Redmon, J., Farhadi, A., 2018. "YOLOv3: An incremental improvement". arXiv:1804.02767

[12] Bochkovskiy, A., Wang, C., Liao, H., 2020. "YOLOv4: Optimal Speed and Accuracy of Object Detection". arXiv:2004.10934