NORDDESIGN CONFERENCE
REYKJAVIK, ICELAND, 12-14 AUGUST 2024

AI Object Detection of Parts in Engineering Drawings and a Concept of Integration in Engineering Education - AssistME

Jan Lennart Kraske, Abdullah Doksanbir, Fabian Dillenhöfer, Bernd Künne

Technical University Dortmund, Germany

Abstract: This article shows how AI and VR technology can be implemented into engineering design education. AI is used to recognize parts in engineering design drawings. The recognition task is done by the Scaled-YOLOv4 object detection algorithm. Implementing adjustments such as tiling and post-logic adjustments regarding the predictions, the detection quality can be improved up to a mAP value of 0.909. The detection of the AI can be transferred into a VR environment via a JSON file. It is shown that the developed AI model is capable of analyzing student's individual solutions.

Keywords: Artificial Intelligence (AI), Object Detection, Virtual Reality (VR), Design Education, Engineering

1 Introduction

AI object detection is used in many different applications such as autonomous driving, education and even in everyday life (Adiguzel et al., 2023; Atakishiyev et al., 2021; Dixit et al., 2022). With the release of ChatGPT, many people were exposed to this technology for the first time and the interest in AI has been increasing ever since. It is also applied to different engineering tasks, for example for analyzing technical drawings. With the help of AI based object detection algorithms, it is possible to extract letters, numbers and symbols out of these images (Eylan et al., 2020, Sarkar et al. 2022) or classify certain objects inside the drawings (Dillenhöfer, 2023). This work demonstrates the AI-based detection of parts in engineering drawings and introduces a concept for its integration into engineering education in section 2.

Many engineering students, especially those in the earlier semesters, have issues imagining technical drawings in three dimensional space. To be a successful engineer it is essential to have well developed spatial skills (Duffy et al., 2018). Spatial understanding of technical drawings requires experience, which many students lack. As a result, many avoidable mistakes are made in the design of mechanical devices such as gearboxes. In the particular system that is presented in this work, AI and virtual reality (VR) technology are combined to extract individual design solutions from technical drawings to visualize and gamify it in a VR environment. Han showed, that the use of VR has a positive influence on learning (Han et al., 2023).

The aim of this work is to investigate whether AI can be used to recognize components in technical drawings with sufficient accuracy to cover the wide variety of student design solutions. The recognized components need to be processed in such a way that they can be used in a VR environment. The approach intends to show the relation between drawing a 3D object and thereby help students understand their own designs better as well as understanding whole machines and their parts. For this purpose, related works are analyzed and the necessary quality requirements are derived. Later, the detection performance of the trained AI is analyzed and discussed.

2 AI object detection and VR in engineering and education

In this section, the state of the art regarding AI object detection particularly in engineering is being introduced. The impact of new technologies integrated in education are shown and which impact they have on the learning process of students.

2.1 AI object detection - short overview

Image recognition is a sub-area of image processing and its task is to identify an object or a feature (feature map) in an image or video (Chai et al., 2021). More and more AI object detection algorithms were revealed recently. One of the most popular object detection algorithms (ODA) is YOLO (You Only Look Once) (Redmon et al., 2016). Object detection algorithms are divided into two classes. There are one-stage and two-stage algorithms. The RCNN algorithms (Masked, Fast and Faster) are examples of two-stage algorithms (Girshick et al., 2016). They differ from one-stage algorithms in the specific focus on regions of interest (ROI). These are predicted by the convolutional neural net and then the objects in these ROIs are labelled in the form of bounding boxes (BB). A BB is an area that reflects the size and position of the predicted object. With single-stage algorithms, the separate creation of ROIs is skipped and the prediction of BB is performed directly in the overall image. As a result, single-stage algorithms are usually faster, but in some cases do not achieve the accuracy of two-stage algorithms. How the predictions of an ODA is categorized can be seen in table 1. If the prediction is right, the BB is classified as true positive (TP). In the second case, there is a BB but not the correct object, making it a false positive (FP) prediction, and if the ODA doesn't detect an object, it's a false negative (FN).

Table 1: Confusion matrix according to (Fawcett, 2006)

result		Actual	
		positive	negative
ODA	positive	True positive (correct detection)	False positive (incorrect marking by ODA)
	negative	False negative (no marking by ODA)	True negative (no detection)

Based on Table 1, quality parameter can be defined as shown in Table 2. Another object detection metric is the mean average precision (mAP) value. To calculate this value, the precision at different recall levels is taken and then averaged over the different classes (Padilla et al, 2020).

Table 2: Quality parameter according to (Fawcett, 2006)

Quality parameter	**Formula**
Missing rate	$\mathrm{MR} = \frac{\mathrm{FN}}{\mathrm{TP} + \mathrm{FN}}$
Precision	$\mathrm{Pre} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$
Recall	$\mathrm{Rec} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{F}N}$
Overall accuracy	$\mathrm{OA} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP}}$

2.2 AI object detection in engineering

After the short overview of ODAs this chapter shows related work. Different ODAs are presented which try to extract information of engineering drawings. With this information it is possible to create a list of requirements the ODA has to fulfill to be suitable for the given task.

Eylan et al. (Eylan et al., 2020) implement AI object detection to recognize parts in hydraulic systems (see Figure 1). For this purpose, the ODA YOLO is being used. This algorithm is able to find trained objects in images and not only classify them but also indicate a region. This box may be a rectangular or, in newer versions, a polygon. Eylan et al. are able to achieve better results with AI than classic image recognition techniques have achieved so far. They used the mentioned detection algorithm to recognize symbols (sensors, valves, etc.) in circuit diagrams (piping and instrumentation diagram). Precision values from 0.67 to 1.0 and accuracies between 0.89 and 1.0 arise. The hydraulic symbols analyzed in that publication are comparable in type and form to the representation of machine elements in technical drawings, or respectively have less developed details. Similar results are shown in Dzhusupova et al. (Dzhusupova et al., 2022). (Eylan et al., 2020)

Researchers at Shanghai Jiao Tong University have investigated object recognition on structural drawings (blueprints). With the help of YOLO, components were recognized and analyzed in a 2D drawing. Information on the position and size of the classes could be extracted. The components are beams, bars and columns. A set of 400 to 1200 images is used as trainings data. The precision of object recognition is between 0.7 and 0.95, with a localization accuracy of 0.1. The authors conclude that there are limitations in the determination of the position and arrangement and the interpretation of the function in the overall system. (Zhao et al., 2020)

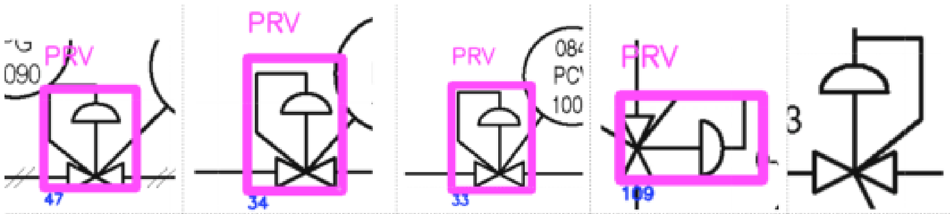

Figure 1: AI object detection in hydraulic drawings with YOLO (Eylan et al., 2020)

Moreno-Garcia et al. worked out a general framework for text and graphics segmentation that is used to identify objects in process and instrumentation diagrams. The framework is capable of digitizing these diagrams. (Moreno-Garcia et al., 2019)

Another use of AI object detection deals with the extraction of text elements in engineering drawings. A company is able to automatically create inspection plans from technical drawings. These can be sampled initially or statistical process control inspection plans are created automatically from 2D CAD, PDF or TIF formats, which can be correctly interpreted from the tolerance and dimensioning information in the technical drawing. Methods and libraries used are not specified. (ELIAS, 2022)

Sarkar et al. published a method on finding and classifying objects in engineering drawings. The object detection is able to classify objects and give their specific position. In total 354 symbols were detected. 301 objects were actually present and 227 were detected correctly. 172 symbols were classified correctly, which results in a precision of 0.4859. (Sarkar et al. 2022)

Another method is proposed by Nguyen et al. (2021). A pattern recognition of technical drawings. Pipeline system are analyzed by detecting objects with specific patterns and performing post-processing task with character recognition to improve the results. (Nguyen et al., 2023)

Faltin et al. put forward a concept to recognize symbols and their orientation in two-dimensional images. Different models are compared that are Keypoint R-CNN, YOLOv7-Pose, and a custom two-stage approach. Keypoint R-CNN, a two-stage-detector, performs best and reaches mAP values up to 0.752. (Faltin et al., 2023).

In the closest related work, Dillenhöfer developed an AI object detection for a part recognition task in engineering design drawings. The system is able to detect objects with a mean overall accuracy of 0.96, which means that 96% of machine components such as bearings are classified true positive. However, there are some limitations in terms of the predicted position that is not examined in detail. The option of automatically generating training data with an algorithm is also examined. The algorithm produces labels for AI object detection training additional to drawings of the shaft and its parts. However, these synthetic generated drawings are not able to improve the detection quality. (Dillenhöfer, 2023; Dillenhöfer et al., 2023) The results of the mentioned related works are summarized in Table 3.

Tabel 3: Overview of related work

	Precision	**Recall**	**mAP0.5**	**OA**
Eylan et al. with yolo	0.67 - 1.0	0.89 - 1.0	-	-
Dzhusupova et al. With yolov4	0.82	0.84	80.6	-
Zhao et al. with yolo	0.853	0.9721	-	0.8331
Sarkar et al. with Faster R-CNN	0.4859	-	-	-
Nguyen et al. with Faster RCNN	-	-	-	0.888
Feltin et al. with Keypoint R-CNN	-	-	0.752	-
Dillenhöfer with yolov4	-	-	-	0.96

2.3 VR and AI in education

Augmented reality (AR) and virtual reality (VR) as teaching tools are used in different fields, such as medicine to teach anatomy (Office of Global Engagement, 2024) or in architecture as a digital design education tool (Aydin and Aktaş, 2020). The learning process can also be improved, for example, by conveying spatial understanding more efficiently. Schnabel and Kvan researched the effectiveness of immersive and non-immersive virtual environments compared to 2d representations of objects. Students were given the assignment to reconstruct a cube with different colored smaller cuboids. Some students were given 2D plans while others used immersive and non-immersive virtual environments. They concluded that while students who used a virtual environment had a lower accuracy when reconstructing the cube compared to those who used 2D plans of the cube, the general understanding of the 3D object could be improved much better with a virtual environment. (Schnabel and Kvan, 2003; Voronina et al, 2019). Doksanbir also published an investigation on the challenges during the learning process and the potential solution with a VR software (Doksanbir et al., 2023). It was shown that engineering students at the TU Dortmund university had issues understanding complex geometry in a technical drawing, which can be explained by the lack of experience to visualize technical drawings as three dimensional objects. Moreover, a VR solution which could improve the students' spatial understanding and motivation to use learning software was welcomed by most students.

Han’s VR application provides an effective engineering education environment and motivates participants to dive deeper into environmental engineering research topics. This is reached by an interactive and immersive experience that helped students to learn about different concepts and environmental challenges. (Han et al. 2023)

Gittinger's systematical review on different studys about spatial abilities and VR shows that the effects of VR on spatial skills show promising results, but need to be further investigated (Gittinger et al., 2023).

Doksanbir pointed out some challenges such as spatial awareness and the understanding of complex parts like the attachment or cast parts that need to be tackled for engineering design education (Doksanbir et al., 2023). The project AssistME verified other problems the students struggle with. These challenges that were pointed out will be addressed by the learning software to improve the competencies that are weak. Hence, different assignments such as an assembly or sorting tasks (recognize parts and categorize them accordingly) are available for the students to attain the competence levels that are needed for design engineers. Hints are being used if the student makes no progress in a given time or if potential mistakes arise in the design. Doksanbir conducted a survey to assess the difficulties of the assignment and gauge the interest of students. The results show among other publications that the spatial understanding is being improved and the interest in a VR solution is present while raising the learning motivation. (Doksanbir et al., 2023)

The area of AI has been attracting attention, especially after the release of ChatGPT. This web interface allows users to type text, which is interpreted and answered by an AI. Despite the increasing popularity, the use of AI in education is still limited. One concept on integrating AI in education is delivered by Garland et al. They integrated AI-vision-systems into a student developed control system. The different student developed projects used the Nvidia Jetson Nano system for image recognition. One example is a robot that monitors chili and determines if these are ripe or unripe. In this way, it is possible to convey knowledge about AI tools to non-experts through experiential methods. (Garland et al., 2023)

3 Concept

The state of the art situation shows that AI based ODA are capable of recognizing parts in engineering drawings. The core research question of this work is if an ODA can be used to recognize and categorize parts in engineering drawings with the necessary accuracy. The endless design solution space could then be covered by this AI.

As shown before, VR is capable of enhancing the spatial abilities. To improve the student’s situation in which they have to learn design by lacking the fundamental skills, this work presents a solution to analyze the student’s drawings and helps students to understand machine parts better by presenting it in an immersive VR. The first step, the analysis of the individual design solution is discussed in this work.

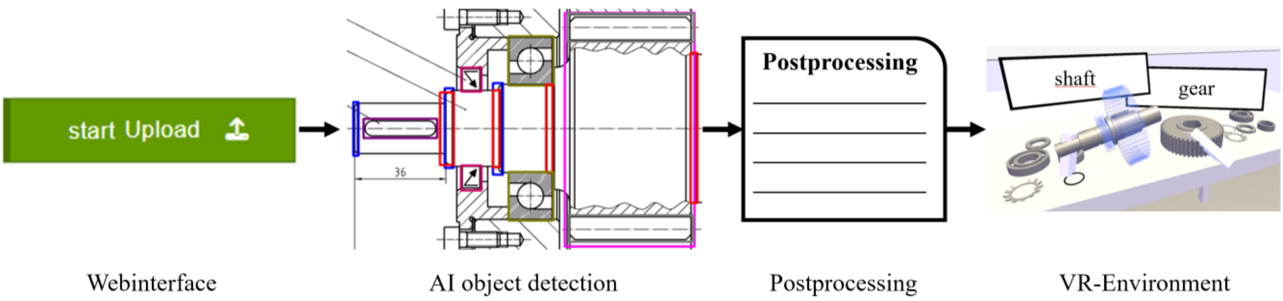

Figure 2: Concept of combining AI and VR technology for engineering design education

The project named AssistME aims at the development of a software that should help engineering students to design machines, such as gearboxes. The students will be able to view their own design in a VR environment and are motivated to complete some tasks such as an assembly of the own construction. The concept was nominated for the XR Awards (Dillenhöfer et al., 2022) in 2022. After showing that assembly tasks in a virtual environment are feasible and the use of VR as a learning medium can increase the spatial skills, the combination of assembly task and a VR is the logical consequence (Dillenhöfer et al., 2022a, Doksanbir et al., 2023, Gittinger et al., 2023). AI ODA is used to cover individual solutions of the students. The simplified concept is displayed in Figure 2. First of all, a website was build and is connected to a server. Here, the students are able to upload their engineering design drawings (technical drawings) and enter some layout design data. After the file was saved on the server, the YOLO AI object detection is triggered via a batch file. The detections are then processed. The detections are compared to lists of standard parts so that the used parts can be identified. Further logical conclusions about the detections are drawn. With this observation process missing or fault, in place or size, detections can be mended. This process creates a text file in JSON format which is provided on the website interface as a download. The text file can then be uploaded into the VR environment in order to depict the individual construction inside it.

The focus of the project is to provide innovative learning material based on the requirements for students. However, the greatest challenge is to develop an object detection that is able to recognize, classify and give the position of the targets reliably. Herein, the algorithm should recognize the following listed parts:

- angular contact ball bearing
- ball bearing
- borehole
- chamfer
- circlip
- feather key
- locking plate groove
- locking plate with shaft nut
- o-ring
- pinion shaft
- radial shaft seal
- relief cut
- shaft nut
- spacer
- tooth
- thread

The state of the art shows that ODAs can be used to detect different machine parts in technical drawings. The accuracy of the related work is between 0.75 and 0.8 in the matter of mAP values and 0.83 and 0.96 in OA. To create an ODA that can recognize the listed objects in the student's drawings a mAP value greater than 0.8 is the goal. The OA has to be close to Dillenhöfer's 0.96. One challenge to achieve this goal is obtaining suitable training data. How this is done is shown in the following.

4 Results

Engineering design drawings include components that can be individual or standard parts. A cutout drawing is shown in Figure 3. Standard parts such as bearings, radial shaft seals, shaft nuts and so on are contained. But also custom designed housings of the gearbox are usually comprised in design drawings of machines. The components differ in terms of detail and size. For example, some kinds of bearings may be difficult to distinguish, because the difference consists only in the outer ring and is very little, it is even difficult to perceive with the human eye. An algorithm automatically generates the image shown in Figure 3 and additionally produces text files with the fitting label that can be used immediately for AI training without any adjustments. Dillenhöfer shows that these synthetic produced images can be utilized for training purpose, but real design drawings would provide better results and that is why these synthetic drawings would be only beneficial if there is a shortage of training material (Dillenhöfer, 2023).

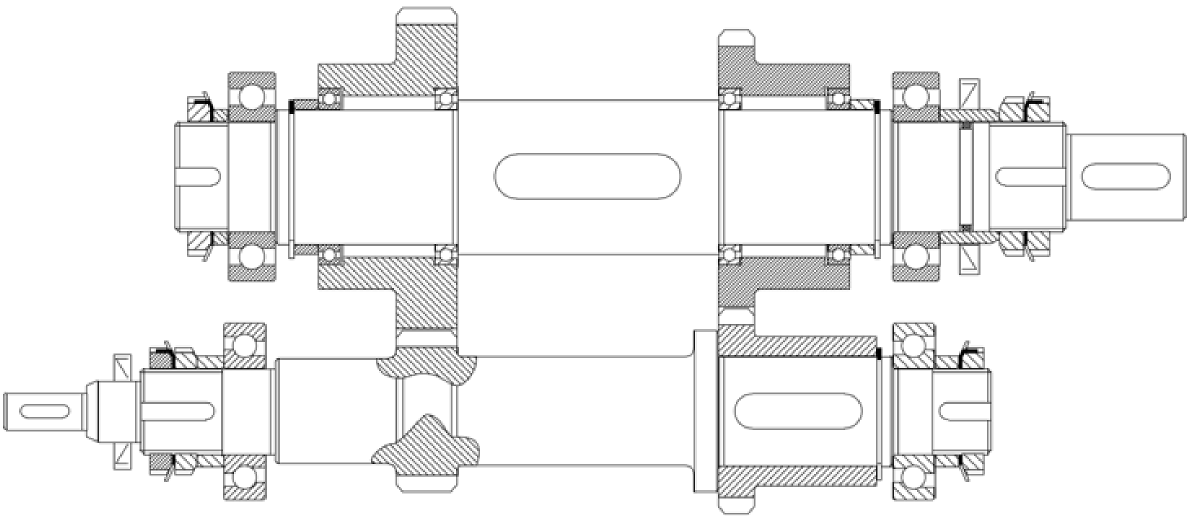

Figure 3: Cutout of an engineering design drawing of a gearbox (namely two shafts) that was generated automatically

When choosing between yolo and rcnn, Scaled Yolov4 was selected for this project as it was the most accurate and scientifically proven algorithm at the beginning of the project (Wang et al., 2021). This ODA was chosen for this project because it was the most up-to-date and scientifically proven version. Papers have already been published that use more recent versions of Yolo, but the versions themselves have not undergone scientific qualification.

4.1 Data procurement

For the training of every AI ODA training data is essential. As Dillenhöfer shows, training data labled drawings should be used (Dillenhöfer, 2023). Therefore, the mentioned objects get manually labled in student's drawings. In each image, about five to 50 labels are marked by hand with the help of a special label software. Only 151 drawings are available for this process. So this work faces two problems with the available data. Firstly, the number of training data is very small and secondly, the drawings are very large. The image tiling approach is used to address this challenge. Processing engineering design drawings using AI ODAs poses a challenge for both software and hardware due to the size of the images. In this case, the images are up to 9362x6623 pixels in size. It is impossible to process images of this size with conventional hardware (in this case an NVIDIA RTX 3090ti). For modelling with Scaled-YOLOv4 it is only possible to process images of sizes up to 3000x3000 pixels with a batch size of one. Since small details should also be recognized in the drawings, it is possible to lose this ability through image compression processes that underlines the ODA. Ünel et al. show that splitting the image into smaller sections leads to an improvement in accuracy (Ünel et al., 2019). This is due to various advantages of tiling. On the one hand, the images are not compressed, which means that no information is lost, and on the other hand, small images have a positive effect on training because less data has to be processed and, thus, the batch size can be increased. The image sections overlap each other by 50% to ensure that no objects are skipped during drawing file generation. Very few objects are larger than 800 pixels and therefore fall out of the grid, as each component is labelled in its sectional view. Another advantage is that the objects are no longer placed exclusively in the center of the image, where the shafts are usually located in the drawings. The ability to move the image section which is being viewed means that the objects appear in different places in the image (see Figure 4). This is a heat map of the objects center point. Red dots show areas where the objects appear more often in the given data. The distribution of objects is very well improved by tiling the images. By tiling the images 2248 tiles are created. So out of 151 images the data amount could be increased to almost 15-fold. Every tile contains at least one object, otherwise its rejected.

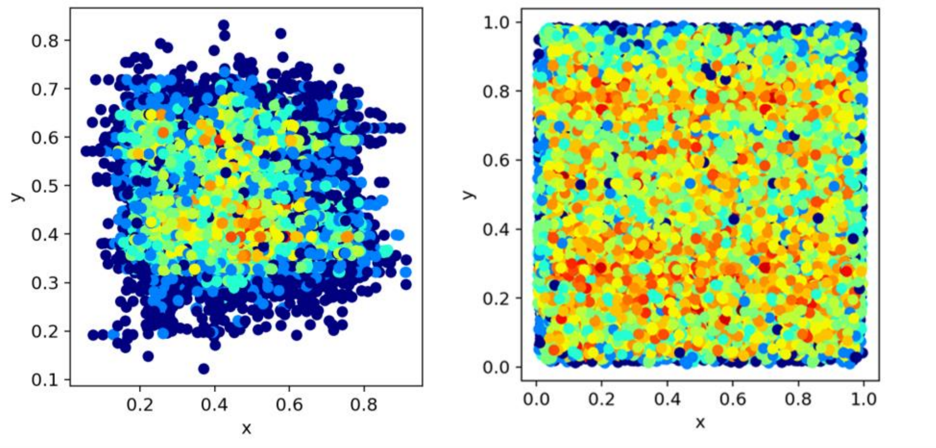

Figure 4: Full range object placing for training without and with tiling

Figure 5 displays the predictions of the ODA that are given by a bounding box, a rectangular marking. The different colors belong to the parts or elements mentioned under the image. In this case, all trained objects have been found so that the classification meets a true positive value of 100%. While the classification reaches mAP values of 0.909, the localization of the objects seems to be worse. For example, the bounding box of the key on the largest diameter of the shaft is larger than it actually is. The precision value of 0.6675 shows that the size of the BB can be further improved. A recall value of 0.97 shows that the given example in figure 5 is not a positive outlier, but the normal case.

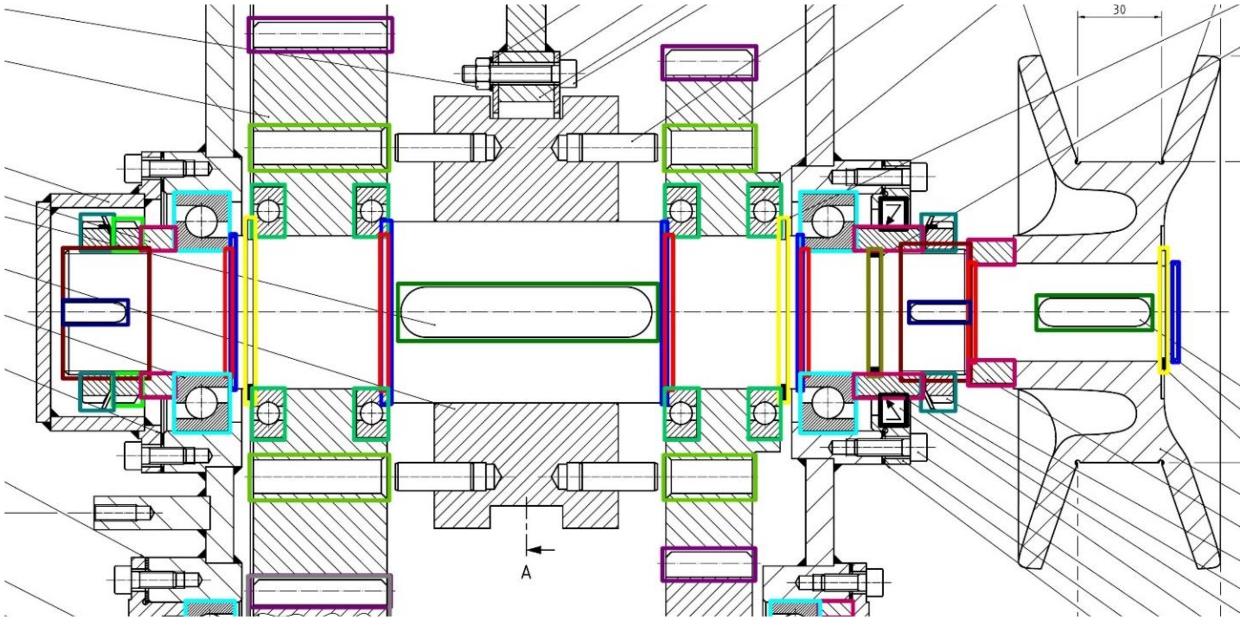

Figure 5: One shaft of engineering design drawing with ODA predictions by Scaled-YOLOv4

4.2 Post-processing logic

In order to optimize the detection quality, a subsequent logic check is executed to diminish the influence of false positive and false negative detections accordingly. Firstly, the detections are sorted. The center lines of the two existing shafts are determined using the centers of characteristic features such as chamfers, relief cuts and keyways. Based on these features, the remaining detections can be assigned to the corresponding shafts. Detections that do not match these shafts are discarded. In this case, a sectional view of the drawing can be seen, which is why a ball bearing is shown above and below the shaft and therefore there may be two markings for one component. The coordinates are calculated for each class using the outer object coordinates and combine those detections. Also, a ratio between pixels and millimeters is formed based on the shaft distance specified by the user. The user input in millimeters is then related to the y-coordinates defined as center lines. Using this ratio, the recognized bearings can be compared with a list of usable bearings in order to read the correct standard parts from the drawing. These can then be output in a parts list and saved in the text transfer file.

Due to the tiling process, objects from classes with a very low confidence value, such as chamfers and relief cuts, may only be partially marked. In order to be able to use these incomplete detections to determine the shaft geometry, these partial markings are enlarged up to the corresponding center line and mirrored on the other side of the center line. In this way, the partial markings are completed and can now contain more information about the shape of the shaft.

Further logical steps are required to be able to infer the geometry of the shafts from the predictions of the AI. By tiling and improving the BB, the detections are so precise that all predictions can be considered in order to draw conclusions about the shaft geometry. To determine the shaft geometry, the predictions are sorted according to their horizontal coordinates and the length of the markings in the vertical direction. Exceptions are feather keys, circlips and gears. For attached components, the inside diameter is considered. If the diameter differs by more than 3 mm, a new section of the shaft is present. Depending on the elements present, the diameter is rounded to the mean value or, in the case of bearings, to the nearest value divisible by five. In the special case of a spacer with an attached radial shaft seal over an o-ring, the inner diameter of the spacer is assumed and the inner diameter of the radial shaft seal is compared with the thickness of the spacer and the shaft diameter. Thanks to the precise predictions of the AI, this method for determining the shaft geometry works very effectively.

4.3 Environmental setting of the ODA

The link between the server website and the VR environment considers user inputs as well as the object detection results. The user input is a prerequisite for using the developed logic to analyze the detections. In this, the students specify which housing design they have selected in their drawing, as well as the center distance, the number of teeth, module and which type of bearing is present for each shaft. This input is required to get the dependence between length in mm and the image size in pixels. Information on the housing type and the gears is transferred to the output file. In addition to the previously collected information, this text file also contains the information obtained by the evaluation logic. Each shaft section is described by length and diameter. Features such as relief cuts, chamfers or feather keys are also included. All attached standard parts are also entered in this file. This file then contains all the necessary information required in the VR environment to provide the required components. The standard parts are then loaded from a library and displayed inside the VR environment (see Figure 6). Inside this environment one can see a shaft in the center and other parts on a table. In this case an assembly task is performed. The gear wheel that is grabbed should be mounted on the shaft in the green highlighted spot that is a help for the students. Furthermore, another blue colored hint is given, where the next part, here a ball bearing, should be mounted next. Figure 6 shows a screenshot of a beta version that is still being developed. The texture and the overall environment will be optimized. The VR controller displayed as sticks will be illustrated as hands.

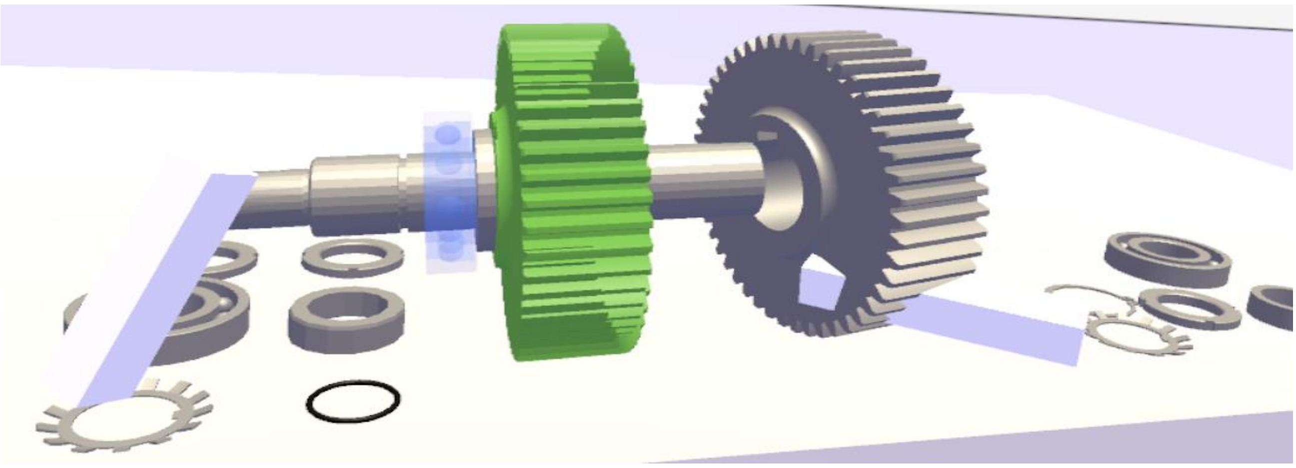

Figure 6: VR environment with assembly task

4.4 Application potential

At some universities, there still exist assignments to design a gearbox by hand. Therefore, the trained model is tested with technical hand drawings and its performance is analyzed. First tests show that the ODA is capable of recognizing the objects with a true positive rate of 55%. However, the detection quality is significantly influenced by both the image quality and how decent the hand drawing was created. Additionally, short tests with regards to a model that was trained with drawings by hand indicated that there is much potential for a better recognition quality.

5 Discussion

The ODA Scaled-YOLOv4 is being used for the recognition task. Newer YOLO versions exist and could provide minor improvements in both accuracy and calculation time. This implies a change in software, libraries and maybe the renewal of label data. Because a mAP value of around 0.909 is reached, only marginal improvements on the object detection quality are expected. The main argument to switch to a newer version of YOLO would be that it is capable of recognizing polygons and is not bounded to rectangular BB. This would mean that all label data need to be renewed, since the rectangular boxes have to be adapted. The automation of the transformation process is not possible. An algorithm could convert only label from polygonal to rectangular shapes.

For the object detection process, 151 drawings were available to train the model. Compared to other applications this is a small number since the literature suggests more training data. With the used tiling method, it was possible to increase the number to 2248 pictures. Nonetheless, the achieved quality values show that this amount of data is enough, even though the increase of training data would only bring minor improvements. In comparison to related work, the developed ODA performs significantly better. Although there is still room for improvement in terms of precision, the average accuracy is better than in the compared works. An overview is given in table 4.

Tabel 4: ODA comparison

	precision	recall	mAP	OA
related work:	0.48 - 1	0.84 - 1	0.75 - 0.8	0.96
AssistME model:	0.6675	0.9708	0.909	-

The project AssistME aims to help students learn machine parts and improve their spatial skills. The performance of the model presented is on such a high level that it can be used as a tool for the analysis of students' individual solutions. The processing of the detections reduces the number of faults and translates the detections into standard parts which then can be added from a library into the VR. With this being said the first step for the realization of the AssistME project is done. In addition, the recognition of hand drawings seems to be possible. The next step will be an evaluation with students, who also helped developing the software, to examine how effectively it conveys the skills that are needed for the engineering education course. Furthermore, it is important to attain feedback from the students, if the software is able to help through the learning process and if it is a useful tool to learn with. The evaluation will take place in the next term, when students need to go through an assignment that aims for an individual designed gearbox. The students will be split into two groups. One group will only take part in the consultation hours while the other group will also have the opportunity to use the software to visualize their design.

The presented work is not restricted to education only, but could also be applied to industrial tasks. For example, if a company needs to digitize drawings, the ODA combined with text recognition algorithms would be capable of such task. Another application would represent the recognition of production error or defects in machines during production or life cycle to predict failure.

Still, this work does meet some limitations. The major part of a machine may be analyzed through the recognition of parts, but not all components can be detected, because custom designed parts like housings are difficult to integrate into this model. Because of this, the learning software works with prefabricated housings. The evolved post logic system is strongly depended on reliable detections. This makes the determination of sizes such as the diameters of the shaft quite difficult. Some of the logic procedure is also dependent on the user input. Moreover, tests with drawings from other applications have revealed that the detection quality could be diminished.

6 Conclusion

Various publications show that it is possible to recognize objects with different YOLO versions. This work adapts the aforementioned ability for a complex scenario as engineering design drawings contain several parts, some of which are very detailed and small. The trained Scaled-YOLOv4 algorithm performs above average with a mAP value of 0.909. This is achieved by using the tiling method to split huge images that are difficult to process into smaller ones. By overlapping the smaller image sections, an even distribution of objects in the images is created. The post-processing logics for evaluating the detections discard unrecognizable detections and can correct incomplete detections. The various techniques that enhance the used YOLO algorithm make a very accurate detection and classification possible. However, there are some limitations that underlie this work. The exact positioning of the objects marked by rectangular BB is not perfectly possible. On the one hand, recognition of the position was not optimized and, on the other hand, rectangular markings are not able to follow custom designs such as circles or curves. In order to achieve better results for localization purpose, a newer YOLO version should be considered, since the recognition of custom formed parts may be possible through the use of polygonal BB. In order to be able to substitute the input without replacement, it is conceivable to recognize various lines with the help of the Hough-transformation to determine the length/image pixel ratio, or to check the evaluation of the shaft geometry.

This work shows that AI can be used to analyze individual drawings. With the help of a VR environment, it will be possible for students to view their individual designs in a playful way. The developed learning software will soon be finished and then provided for education.

Acknowledgement

We gratefully acknowledge financial support for AssistME through the foundation “Innovation in der Hochschullehre” (Innovation in higher education) with funds pro-vided by the Toepfer Stiftung gGmbH (Foundation) under Grant No. FRFMM-740/2022.

References

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. Contemporary Educational Technology, 15(3), ep429. https://doi.org/10.30935/cedtech/13152.

Atakishiyev, S., Salameh, M., Yao, H., Goebel, R. (2021). Explainable artificial intelligence for autonomous driving: A comprehensive overview and field guide for future research directions. arXiv preprint arXiv:2112.11561.

Aydin, S., Aktaş, B. (2020): Developing an Integrated VR Infrastructure in Architectural Design Education. Frontiers in robotics and AI, 7, 495468. https://doi.org/10.3389/frobt.2020.495468.

Chai, Junyi; Zeng, Hao et al. (2021): Deep learning in computer vision: A critical review of emerging techniques and application scenarios, Machine Learning with Applications Volume 6, ISSN 2666-8270, https://doi.org/10.1016/j.mlwa.2021.100134.

Dillenhöfer, F., Doksanbir, A., Künne, B. (2022): AssistME, Support tool for aspiring design engineers, In: XR Science Awards, nominated for ther best concept; https://divr.de/award-en/.

Dillenhöfer, F., Künne, B., & Willms, U. (2022a). Neologised teaching concept and materials combine remote teaching and hands-on activities. *2022 IEEE German Education Conference (GeCon)*, Article 9942764. Publiziert. GECON, Berlin. https://doi.org/10.1109/gecon55699.2022.9942764.
Dillenhöfer, Fabian (2023): Entwicklung einer KI-Objekterkennung für technische Zeichnungen im Maschinenbau (engl.: Development of an AI object detection for technical drawings in mechanical engineering), Dissertation, TU Dortmund University, http://dx.doi.org/10.17877/DE290R-24140.
Dillenhöfer, F.; Doksanbir, A., Kraske, J. L.; Künne, B. (2023): Automatic Generation of Training Data for AI Object Detection in Terms of Technical Drawings in Engineering. In M. E. Auer, R. Langmann, & T. Tsiatsos (Hrsg.), *Open science in engineering* (Verlagsversion, 1. Aufl., Bd. 763, S. 643–651). Springer Nature Switzerland; https://doi.org/10.1007/978-3-031-42467-0_60.
Dixit, P., Bhattacharya, P., Tanwar, S.,, Gupta, R. (2022). Anomaly detection in autonomous electric vehicles using AI techniques: A comprehensive survey. Expert Systems, 39(5), e12754.
Doksanbir, A., Dillenhöfer, F., & Künne, B. (2023). Work-in-progress: a study on the problems of engineering students designing gearboxes and VR as a possible solution. In M. E. Auer, R. Langmann, & T. Tsiatsos (Hrsg.), *Open science in engineering* (Verlagsversion, 1. Aufl., Bd. 763, S. 527–534). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42467-0_48.
Duffy, G., Sorby, S., Reves, P.R., Delahunty, T., Perez, L. and Ravishankar, J. "The Link between Spatial Skills and Engineering Problem-Solving," *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, Wollongong, NSW, Australia, 2018, pp. 272-278, doi: 10.1109/TALE.2018.8615193.
Dzhusupova, R., Banotra, R., Bosch, J., Olsson, H. (2022): Pattern Recognition Method for Detecting Engineering Errors on Technical Drawings. In: 2022 IEEE World AI IoT Congress, AI-IoT 2022, pp. 642-648. 10.1109/AIIoT54504.2022.9817294.
ELIAS GmbH (2022): https://www.elias-gmbh.de/, last accessed: 10.02.2024.
Elyan, E.; Jamieson, L.; Ali-Gombe, A. (2020): Deep learning for symbols detection and classification in engineering drawings. Neural networks the official journal of the International Neural Network Society 129, S. 91–102.
Faltin, Benedikt; Schönfelder, Phillip; König, Markus. (2023). Improving Symbol Detection on Engineering Drawings Using a Keypoint-Based Deep Learning Approach.
Fawcett, Tom (2006): An introduction to ROC analysis. In: Pattern Recognition Letters 27 (8), S. 861–874. DOI: 10.1016/j.patrec.2005.10.010.
Garland, Nigel Patrick; Wade, Russell; Palmer, Sarah (2023): NON-EXPERT PRACTICAL APPLICATION OF AI VISION SYSTEMS IN DESIGN ENGINEERING PROJECTS. In: Proceedings of the International Conference on Engineering and Product Design Education, EPDE 2023. 25th International Conference on Engineering and Product Design Education, 7th and 8th September 2023: The Design Society.
Gittinger, M., & Wiesche, D. (2023). Systematic review of spatial abilities and virtual reality: The role of interaction. *Journal of Engineering Education*, 1–20. https://doi.org/10.1002/jee.20568
Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2016). "Region-Based Convolutional Networks for Accurate Object Detection and Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 142-158, 1 Jan. 2016, doi: 10.1109/TPAMI.2015.2437384
Han, B., Weeks, D.J. and Leite, F. (2023), Virtual reality-facilitated engineering education: A case study on sustainable systems knowledge, *Comput. Appl. Eng. Educ*. 2023; 31: 1174–1189. https://doi.org/10.1002/cae.22632
Moreno-García, C.F., Elyan, E., Jayne, C. (2019): New trends on digitisation of complex engineering drawings. *Neural Comput & Applic* **31**, 1695–1712; https://doi.org/10.1007/s00521-018-3583-1.
Nguyen, T., Pham, L. V.; Nguyen, C.; Nguyen, V. V (2023). Object Detection and Text Recognition in Large-scale Technical Drawings. Proceedings of the 10th International Conference on Pattern Recognition Applications and Methods. https://doi.org/10.5220/0010314406120619.
Office of Global Engagement, Taipei Medical University (2024), https://oge.tmu.edu.tw/tmu-pioneers-worlds-largest-virtual-reality-anatomy-class/, last accessed 2024/02/02.
Padilla, R., Netto, S. L., da Silva, E. A. B. (2020), "A Survey on Performance Metrics for Object-Detection Algorithms," *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Niteroi, Brazil, 2020, pp. 237-242, doi: 10.1109/IWSSIP48289.2020.9145130
Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779–788). https://doi.org/10.1109/CVPR.2016.91
Schnabel, M. A., Kvan, T. (2003): Spatial Understanding in Immersive Virtual Environments. In-ternational Journal of Architectural Computing vol. 1.
Sarkar, Sourish; Pandey, Pranav; Kar, Sibsambhu (2022): Automatic Detection and Classification of Symbols in Engineering Drawings, https://arxiv.org/abs/2204.13277.
Ünel, F. Ö. ; Özkalayci, B. O.; Çiğla, C. (2019): "The Power of Tiling for Small Object Detection," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Long Beach, CA, USA, 2019, pp. 582-591, doi: 10.1109/CVPRW.2019.00084. .
keywords: {Object detection;Real-time systems;Proposals;Feature extraction;Task analysis;Agriculture;Image resolution},
Voronina, M., Tretyakova, Z., Krivonozhkina, E., Buslaev, S., Sidorenko, G. (2019).: Augmented Reality in Teaching Descriptive Geometry, Engineering and Computer Graphics – Systemat-ic Review and Results of the Russian Teachers' Experience. Eurasia Journal of Mathematics, Science and Technology Education. 15. 10.29333/ejmste/113503.
Wang, C.-Y., Bochkovskiy, A. & Liao, H.-Y. M. (2021). Scaled-YOLOv4: Scaling Cross Stage Partial Network.. *CVPR* (p./pp. 13029-13038), : Computer Vision Foundation / IEEE.
Zhao, Y.; Deng, X.; Lai, H. (2020): A Deep Learning-Based Method to Detect Components from Scanned Structural Drawings for Reconstructing 3D Models. Applied Sciences 6/10, S. 2066.

Contact: Jan Lennart Kraske, TU Dortmund, Department of Machine Elements, Leonhard-Euler-Straße 5, 44135, Dortmund, Germany, +492317555734, +492317552740, Lennart.kraske@tu-dortmund.de