A Simple Screw-Hole Discrimination Pipeline for Deployment in Autonomous Manufacturing

1st Nils Mandischer
IGMR
RWTH Aachen University
Aachen, Germany
mandischer@igmr.rwth-aachen.de

2nd Sebastian Döbler
IGMR
RWTH Aachen University
Aachen, Germany
doebler@igmr.rwth-aachen.de

3rd Burkhard Corves
IGMR
RWTH Aachen University
Aachen, Germany
corves@igmr.rwth-aachen.de

***Abstract*—Autonomy in manufacturing is a major challenge of Industry 4.0. To allow autonomous assembly of rotary tables, one important aspect is the localization and discrimination of empty holes and screws. In this paper, we propose a novel feature descriptor to discriminate screws and holes. The descriptor is embedded in an end-to-end pipeline, which localizes candidates on concentric workpieces. The candidates are analyzed using a slim tree-based classifier, which discriminates inserted screws and empty holes. These are later used in autonomous screwing processes. The proposed methods are supplemented by an intuitive UI-driven process, which enables non-expert users to train and deploy the them leniently in autonomous manufacturing processes. The feasibility of the proposed method is demonstrated in multiple lab experiments and in a real-world demonstrator as part of the SHAREWORK project.**

***Index Terms*—Screw Detection, Autonomous Manufacturing, Explainable AI, Random Forest**

I. Introduction

Aim of the SHAREWORK project is to bring human-robot collaboration (HRC) to less automated industries and to make the framework's methods easily accessible to non-expert users. In collaborative processes, the robot is required to identify certain objects within its environment. One important task is to discern whether screws are inserted into a rotary table. Based on the detection, the manufacturing process of the rotary table is supervised, and the robot enabled to react dynamically to human actions while assisting the assembly process.

In this paper, we propose a simple discrimination pipeline[1] for screw and hole identification to make AI accessible to the non-expert user. First, candidates are identified based on geometric segmentation. These are, then, analyzed in a binary classifier, which discriminates inserted screws from holes. The method is validated in lab tests and in the real application. Latter shows the feasibility of the technology to be integrated by non-experts. The main contributions of this paper are:

- Classifier to discriminate overlying or countersunk screws and empty holes using a novel feature descriptor.
- End-to-end pipeline to be used on concentric workpieces.
- Guided training and deployment approach to ease use of the proposed pipeline for non-expert users.

We like to thank the European Union for funding this work within the SHAREWORK project (grant 820807).

[1]Code available at https://github.com/nilsmandischer/sharework_screw_detection (add "_ros" for ROS interface).

II. Related Work

The reliable and resource efficient detection of objects is one of the fundamental tasks in autonomous manufacturing and HRC. Ramana, Choi, and Cha [1] and Wegener et al. [2] both use Cascade classifiers trained using AdaBoost to detect screws in an image. Ramana et al. additionally use a Support Vector Machine (SVM) to classify the located screws into different groups. Bdiwi, Rashid, and Putz [3] propose a method to detect screws in an RGBD scan. Their approach first uses a Harris Corner Detector to find screw candidates with a high False Positives (FP) rate and filtering them, subsequently. Filtering consists of an analysis of the region of interest's (ROI) hue, saturation, and brightness values, and the usage of the camera's depth measurements. Cruz-Ramirez et al. [4] suggest the usage of multi template matching to detect screws in an HRC environment to dismantle metal-ceiling structures in buildings. This approach is applicable, if a single type of screw is used, but does not translate well to scenarios with different types of screws, since a new template is needed for each type. Li, Wei, and Xing [5] propose a similar method, which uses local binary patterns and a sliding window approach to localize screws and generate features which are, then, classified using a SVM. Similarly, Tellaeche, Maurtua, and Ibarguren [6] propose a 2D matching method using 3D CAD models to match onto an image.

Tree classifiers, like Random Forest Classifiers (RFCs), may also be used for object classification [7], but do not see widespread use for detecting screws in images. Another approach for locating and classifying objects is the usage of neural networks [8]. Depending on the specific implementation, before classifying screws they first have to be located. Yildiz and Wörgötter [9] and Cha, You, and Choi [10] use the Hough Circle Transform (HCT) to find the location of possible screw candidates and classify them using a neural network. Martinez, Ahmad, and Al-Hussein [11] propose an ellipse fitting algorithm - instead of HCT - to locate circular objects, like screws, to compensate for different observation angles.

Since it is the scope of SHAREWORK to enable untrained personnel to use all methodology provided by the project's framework, it is unreasonable to apply deep learning techniques, as they are not able to adopt the methodology, accordingly. The same applies to any technique which requires specifically trained personnel to retrain a model or generate templates for matching. Instead, we propose a heuristic feature descriptor in combination with a RFC and an intuitive user interface (UI) for retraining the classifier. The benefit is that the user can provide new training data to adapt to new models of the rotary table or screws, and train the classifier by themselves without knowledge of the underlying methodology.

III. Methodology

The methodology is split into three stages: (A) locate and segment screw/hole candidates, (B) extract features, (C) classify candidates into screws and holes (see Figure 1). First, candidates are found on the target area. They are analyzed for certain features, which typically characterize inserted screws and holes. Finally, the features are fed in a classifier that decides to which class the candidates belong.

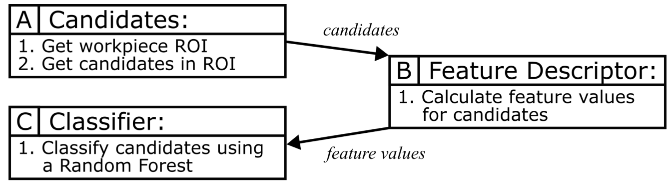

Fig. 1: Sketch of the end-to-end pipeline.

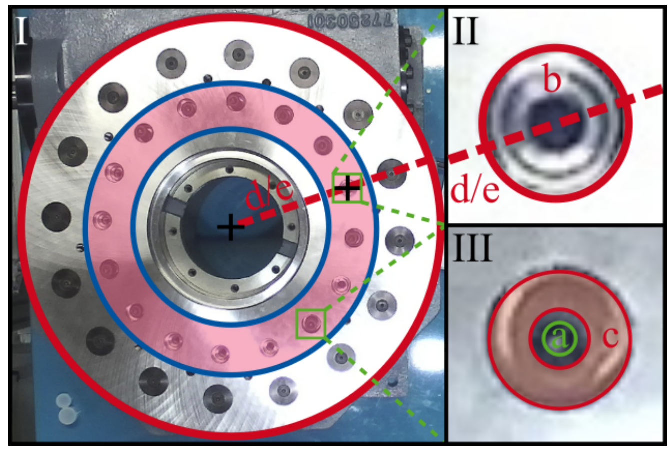

Fig. 2: *I*: Detected rotary table (red circle) and corresponding ROI (red area in blue lines). Two candidates (green) and the line used in features (d) and (e) for one of the candidates (red line). *II*: Detected candidate (red circle) and the line used in features (d) and (e) (red line); depicted is a hole. *III*: Detected candidate and the areas for feature (a) (green) and (c) (red); depicted is a screw.

A. Candidates

To locate candidates for classification, we exploit geometric features of the rotary table. The holes are distributed concentric on a slim circle segment of the rotary disk (see Figure 2.I). First, the rotary disk (red circle) is detected using a HCT. Since the diameter of the workpiece is known, the ROI (red area) in which the candidates are located is well defined and is determined accordingly. To identify the candidates, HCT is used on the previously determined ROI. This allows for reliable detection of only the desired candidates, hence makes classification easier, as there is a highly controlled input of only two distinct cases, holes with and without screws. The proposed approach to locate candidates can be directly used for any circular workpiece with concentrically aligned screws or adapted for arbitrarily shaped workpieces with predefined screw locations. Note, that the features descriptor and classifier do not require the workpiece to be concentric. However, the approach to determine the center axis in Section III-B may need to be adapted for non-concentric workpieces.

B. Feature Descriptor

The selected features follow a minimalist approach to save computational time for training the model and during subsequent execution, and to allow better explainability for the user. Hence, the number of features is kept as small and understandable as possible, while producing reliable results. All features are calculated using a gray-scale representation of the image. The ROI of a candidate is the area inside the detected circle of the HCT, which is represented by a red circle in Figure 2.II. The usage of gray-scale images results in the actual screw color being negligible and the difference in gray-scale between the screw and the workpiece being more important. In the industrial use-case, silver and black screws on a metallic surface are used, which are either overlying or countersunk. Black screws can be understood as screws with a strongly different color than the workpiece and silver screws as of similar color. Note, by this the descriptor may be used for all type of colored screws. We describe five features:

(a) Bright Center The feature aims to identify a screw by its characteristic bright center spot given that the work bench is properly illuminated - which is the usual case. The feature describes the average intensity of a small circular area at the center of the detected candidate (see Figure 2.III). The area used to calculate the feature has a diameter of 25% of the candidate's diameter.

(b) Black Screw This feature takes the average intensity of the whole area inside the candidates ROI to account for the usage of black screws, which are darker than the metal of the surrounding parts (see Figure 2.II). This feature does not limit the method to the usage of black screws, but is a key feature in case of screws with a different color than the workpiece.

(c) Bright Metal Holes without a screw inserted show a brighter area in which usually the screw head resides. Therefore, the feature computes the intensity of the outer ring of the candidate (see Figure 2.III). The inner diameter of the ring is 50% and the outer diameter is 100% of the diameter of the detected circle.

(d) Intensity Plateaus For this feature, we determine a line across the candidate which crosses the rotary disk's and the candidate's center point (see Figure 2.I and 2.II). This corresponds to the exit of the incident light. Note, that the workpiece is illuminated concentrically. Along this line, inside the ROI of the candidate, the number of intensity plateaus is counted. An intensity plateau is defined as a sequence of pixels, which are of roughly equal intensity, without being separated by more than four pixels of different intensity. This is to consider the different intensity distributions along the axis of the hole, and is particularly important in case of countersunk screws.

(e) Intensity Length This feature uses the same line as in (d), but computes the number of pixels, which are of greater intensity than the average of the whole ROI of the candidate.

C. Classifier

Typically, a RFC is not the best suited classifier for real-time applications since predictions take longer compared to other classical learning-based classifiers. This is due to the large amount of individual decision trees generated for a Random Forest. This property, however, can be neglected due to the small feature count. The classifier is implemented with OpenCV. For training, the maximum tree depth is set to 25 and the minimum required samples for a node to be split is set to three. The regression accuracy is set to 0.01 and the Random Forest is allowed to build surrogate splits. The maximum categories are set to two and the subset of features for each tree is set to three. The learning process is terminated after 200 iterations. The a priori class probabilities are set to be equal for all classes.

IV. Training and Deployment

As the proposed method uses only a small number of training samples, those are generated by the end-user directly. For this purpose, we provide a simple UI that segments an image and presents the user with the candidates (see Figure 3). The user then decides if a hole, screw, or false detection is depicted by pressing the according button. After all images have been processed the classifier is automatically trained and the application is ready to be used. With this approach, the classifier can be trained in under five minutes including taking images. This measure is taken from the industrial use-case presented later in Section V-B. In the scenario, the hardware integrator was able to apply the UI without further instructions, which underlines the simplicity and intuitivity of the approach. To parameterize the full pipeline, only the diameter of the screws and the rotary table have to be set. Based on the observations, the UI is termed feasible. In future additions, the text-based UI should be embedded in a graphical user interface (GUI) to improve on usability.

V. Validation

To validate the method, two different setups are used: First, the method is validated in a lab scenario as seen in Figure 4 to determine key performance indicators (KPIs). Secondly, the method is integrated in an industrial use-case in cooperation with the Goizper Group as seen in Figure 2.

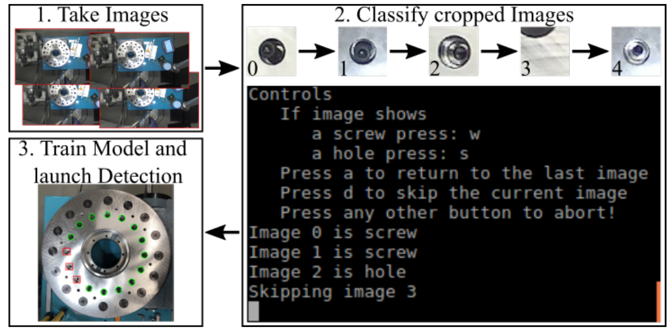

Fig. 3: Simple UI used for training the classifier in the industrial use-case.

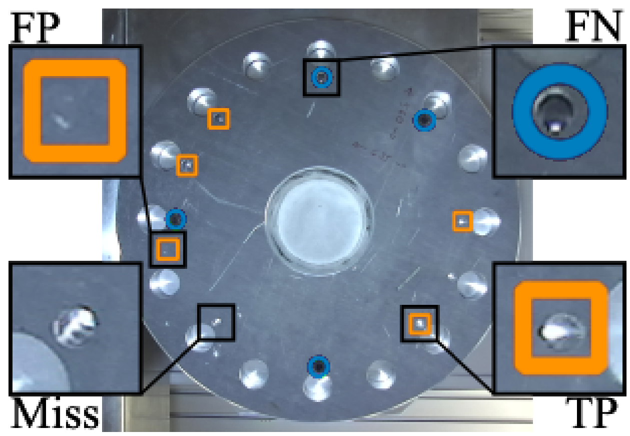

Fig. 4: Exemplary detections on mock (lab scenario): Detected screws (blue circle) and holes (orange square) are highlighted (re-colored for convenience) including malformed output.

A. Lab Scenario

In lab validation, a Stereolabs ZED 2 camera (2208x1242px) is placed at a distance of 0.475m to a mock of the rotary table to detect up to ten overlying M6 screws (see Figure 4). Four sets of images are generated which cover 1000 object samples each and include silver or black screws with varying lighting conditions. Out of the 1000 object samples, 50% are screws and the other 50% are empty holes respectively. To train the classifier, we use samples of a rotary disk with all or no screws inserted in three different orientations, which results in 60 training samples total. The samples are taken with black and silver screws separately. When classifying objects, four distinct outcomes can be achieved (see Figure 4):

TP Candidate found, and classified correctly
FN Candidate found, and classified wrongly
FP Candidate found, but located wrongly
Miss Candidate not found at all

The agglomerated occurrences of True Positives (TP), False Negatives (FN), False Positives (FP), and Misses are listed in Table I for all test-cases of the lab scenario. It is noticeable that screws are less often missclassified or not found at all when compared to holes. Furthermore, almost all falsely segmented candidates are classified as holes, which is expected, as the proposed method does not distinguish objects from background in case of erroneous segmentation.

For validation, we define KPIs for the full pipeline, including segmentation errors (Misses and FP). The indicators are accuracy ac, precision pr, and recall re, defined by

$$ac = \frac{TP}{TP + FN + FP + Miss}, \tag{1a}$$

$$re = \frac{TP}{TP + FN + Miss}, \tag{1b}$$

$$pr = \frac{TP}{TP + FP}. \tag{1c}$$

In addition, as the segmentation may be exchanged in applications with non-concentric workpieces, we also measure the performance without influence of the HCT in form of adjusted accuracy recall $\tilde{re}$, defined by

$$\tilde{re} = \frac{TP}{TP + FN}. \tag{2}$$

To analyze the adaptability of the proposed method, we define five different scenarios in which KPIs are measured:

- **A** Well illuminated workplace with black screws
- **B** Well illuminated workspace with silver screws
- **C** Sparsely illuminated workspace with silver screws
- **D** Subset of set A, trained on sets A and B
- **E** Subset of set B, trained on sets A and B

The subsets and training sets in D and E contain an equal amount of black and silver screws, and for training 60 screw and 30 hole samples are used. Sets A, B, and C use training samples generated with according illumination (well-lit or sparse) and screws (black or silver). The results in each scenario and corresponding KPIs are listed in Table II. In general, all test cases show good performance in detecting screws on the rotary disk. While we define features which benefit illuminated reflective surfaces, they contribute to the overall class decision and do not lead to over-fitting on the lighting conditions, which is particularly emphasized by case C. In case of silver screws, using a sparsely illuminated workspace benefits the classification. This is due to the fact that brightly illuminated screws (case B) loose contrast to the metallic surface of the rotary disk, hence the enlarged number of Misses, FP, and FN compared to case C. In addition, cases D and E show that the classifier may also be trained with screws of different colors and used in more varied applications with diverse workshop items. However, in case of multi-screw training (D, E), the KPIs are slightly impaired compared to single-screw samples (A, B, C). From the lab evaluation, we conclude that the proposed classifier is suited for detecting screws in a workshop scenario with changing lighting conditions (e.g., over the day) and screw types. The lab scenario is further validated in a full manufacturing mock application, which is discussed in [12].

B. Industrial Scenario

The industrial scenario is set up in cooperation with the Goizper Group as part of the SHAREWORK project [13]. In a HRC scenario, a human inserts screws into the rotary disk, the robot detects, and, consequently, uses a powered screw driver to fasten the screws. For detection, the robot is driven in a configuration such that the camera plane is concentric and parallel to the rotary disk (see Figure 2). Compared to the mock in Section V-A, the real rotary table has 16 black M10 countersunk screws. In the industrial scenario, the proposed method is not only used for detection, but also for localization of the position of inserted screws.

To validate the industrial scenario, the same five test cases and KPIs as in the lab scenario are used. Tests use 1500 samples each, the training procedure and training data amount are equal to the lab case. The results and corresponding KPIs are listed in Table III. In general, the results are comparable to the lab scenario if the candidate is correctly segmented. However, the number of undetected or wrongly detected holes is significantly higher. This is due to sub-optimal lighting conditions. In the lab scenario, a spot lamp is mounted directly over the rotary disk and a diffuse light source is placed to its side. The diffuse light has a strong light intensity. In the industrial scenario, the base light is way less intense. Hence, cases A and B are closer to C. In the lab scenario, case C benefits from the better base illumination, hence, the major decline in case C. To improve on the results, an external diffuse light would have to be mounted close to the workpiece, which was not viable in our test case. However, the lower values in the KPIs by no means indicate

TABLE I: Confusion matrix for all lab test cases combined (GT: ground truth).

GT\Estimate	Screw	Hole	None
Screw	1902	93	2
Hole	141	1827	36
None	8	57	-

TABLE II: Total occurrences of TP, FP, Miss and FN over all test sets of the lab scenario and corresponding scores. All measures acc, pr, re, and $\tilde{re}$ are in %.

Set	TP	FP	Miss	FN	acc	pr	re	$\tilde{re}$
A	964	18	16	20	94.7	98.2	96.4	98.0
B	893	23	12	95	87.3	97.5	89.3	90.4
C	999	1	1	0	**99.8**	**99.9**	**99.9**	**100**
D	472	10	5	34	90.6	97.9	92.4	93.3
E	401	13	4	85	<u>79.7</u>	<u>96.9</u>	<u>81.8</u>	<u>82.5</u>
D + E	873	23	9	119	85.3	97.4	87.2	88.0
Total	3729	88	38	234	91.2	97.7	93.2	94.1

TABLE III: Total occurrences of TP, FP, Miss and FN over all test sets of the industrial scenario and corresponding scores. All measures acc, pr, re, and $\tilde{re}$ are in %.

Set	TP	FP	Miss	FN	acc	pr	re	$\tilde{re}$
A	1299	0	6	195	86.6	**100**	86.6	86.9
B	1357	0	0	143	90.4	**100**	90.4	90.4
C	1131	0	7	362	75.4	**100**	75.4	75.7
D	1069	0	0	431	<u>71.2</u>	**100**	<u>71.2</u>	<u>71.3</u>
E	1379	0	4	117	**91.9**	**100**	**91.9**	**92.2**
D + E	2448	0	4	548	81.6	**100**	81.6	81.7
Total	6235	0	17	1248	83.1	100	83.1	83.3

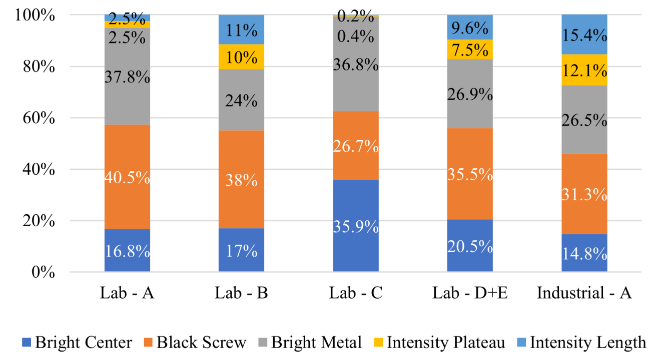

Fig. 5: Importance of features along selected test cases.

that the method is unsuited for the lab scenario. Indeed, during testing[2,3], the robot is able to autonomously fasten the screws successfully [13]. This indicates that the method is sufficiently accurate in both, detection and segmentation (i.e., position estimate of the screws) to insert the powered screw driver into the hexagon socket (compare [14], [15]) and screw them in, accordingly. In the demonstrator application, human and robot are able to successfully and repeatably assemble the rotary disk on the rotary table. Throughout, only minor delays due to erroneous detections are observed, which mostly come from partial occlusion while handing over the workpiece (in addition to the already described challenges). Therefore, we can conclude that the proposed method is robust and accurate enough to be used in autonomous assembly tasks in real-world environments.

C. Feature Descriptor

Finally, besides the general feasibility, the feature descriptor itself is analyzed to give an understanding of its situational suitability. The importance of each feature on the class decision for both scenarios is shown in Figure 5. Noticeably, features *Black Screw* and *Bright Metal* are most important in all cases independent of screw type or illumination. In the lab scenario, certain features dominate, while in the industrial scenario, the feature importance is better distributed among all features. In particular, the *Intensity Pleateau* and *Intensity Length* feature become more important in case of countersunk screws. This indicates that the features in the descriptor are likewise well suited for the generic industrial scenario. However, in certain situations adapting the feature set may benefit the overall performance of the embedding classifier, as, particularly in sparse lighting, certain features may be omitted. This may accelerate the pipeline as these features do not need to be computed.

VI. Conclusion

In this paper, we proposed an end-to-end pipeline to discriminate screws from holes within the application of autonomous manufacturing of concentric workpieces. The proposed methods are designed for overlying or countersunk silver and black hexagon socket screws on a metallic workpiece. To find candidates, we apply the Hough Circle Transform combined with a region of interest that is easily parameterized by the real dimensions of the workpiece. The classifier uses a novel feature descriptor, which is composed of five simple and understandable features. As classifier, a Random Forest is used. For training, we apply an UI-based process to guide non-expert users from plain images to a fully trained classifier, ready for application. The proposed methods and the UI were integrated in five lab scenarios and an industrial scenario. The subsequent validation demonstrated good performance and usability, particularly in the industrial case, which demonstrated a successful human-robot collaborative manufacturing process.

[2]https://sharework-project.eu/integration-of-goizper-mock-up-at-stams-laboratory/

[3]https://www.youtube.com/watch?v=vNPjjKpIDWw

References

[1] L. Ramana, W. Choi, and Y.-J. Cha, "Fully automated vision-based loosened bolt detection using the viola–jones algorithm," *Structural Health Monitoring*, vol. 18, no. 2, pp. 422–434, 2019.

[2] K. Wegener, W. H. Chen, F. Dietrich, K. Dröder, and S. Kara, "Robot assisted disassembly for the recycling of electric vehicle batteries," *Procedia CIRP*, vol. 29, pp. 716–721, 2015.

[3] M. Bdiwi, A. Rashid, and M. Putz, "Autonomous disassembly of electric vehicle motors based on robot cognition," in *International Conference on Robotics and Automation*. IEEE, 2016, pp. 2500–2505.

[4] S. R. Cruz-Ramirez, Y. Mae, T. Takubo, and T. Arai, "Detection of screws on metal-ceiling structures for dismantling tasks in buildings," in *International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 4123–4129.

[5] C. Li, Z. Wei, and J. Xing, "Online inspection system for the automatic detection of bolt defects on a freight train," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 230, no. 4, pp. 1213–1226, 2016.

[6] A. Tellaeche, I. Maurtua, and A. Ibarguren, "Use of machine vision in collaborative robotics: An industrial case," in *21st International Conference on Emerging Technologies and Factory Automation*. IEEE, 2016, pp. 1–6.

[7] Z. Liu and H. Xiong, "Object detection and localization using random forest," in *Second International Conference on Intelligent System Design and Engineering Application*. IEEE, 2012, pp. 1074–1078.

[8] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in *Advances in Neural Information Processing Systems*, vol. 26. Curran Associates, Inc, 2013.

[9] E. Yildiz and F. Wörgötter, "Dcnn-based screw detection for automated disassembly processes," in *International Conference on Signal-Image Technology & Internet-Based Systems*. IEEE, 2019, pp. 187–192.

[10] Y.-J. Cha, K. You, and W. Choi, "Vision-based detection of loosened bolts using the hough transform and support vector machines," *Automation in Construction*, vol. 71, pp. 181–188, 2016.

[11] P. Martinez, R. Ahmad, and M. Al-Hussein, "Real-time visual detection and correction of automatic screw operations in dimpled light-gauge steel framing with pre-drilled pilot holes," *Procedia Manufacturing*, vol. 34, pp. 798–803, 2019.

[12] S.-O. Bezrucav, N. Mandischer, and B. Corves, "Artificial intelligence task planning of cooperating low-cost mobile manipulators: A case study on a fully autonomous manufacturing application," *Procedia Computer Science*, vol. 217, pp. 306–315, 2022.

[13] A. Umbrico, M. Ansagasti, S.-O. Bezrucav, F. Canale, A. Cesta, B. Corves, N. Mandischer, M. Mondragon, C. Naso Rappis, and A. Orlandini, "Enhanced cognition for adaptive human-robot collaboration," in *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2022.

[14] ISO 2936:2014, "Assembly tools for screws and nuts - hexagon socket screw keys," 2014-05-01.

[15] ISO 4762:2004, "Hexagon socket head cap screws," 2004-03-01.