The International Journal of Advanced Manufacturing Technology
https://doi.org/10.1007/s00170-026-18916-7

ORIGINAL ARTICLE

Visual detection of flap tears on abrasive flap wheels using deep learning

Falko Kähler[1] · Jessica Ehrbar[1] · Thorsten Schüppstuhl[1]

Received: 26 May 2026 / Accepted: 6 August 2026

Abstract

During abrasive flap wheel grinding, flaps may fail and tear prematurely. Such flap tears increase local tool flexibility and can impair the uniformity of material removal in automated grinding processes. To ensure process reliability, this paper proposes a novel computer vision approach for the automated detection and quantification of flap tears using deep learning. The processing pipeline is based on images showing the radial tool structure, followed by a geometric normalization that transforms the flap arrangement into a linear coordinate system. A YOLOv8-based object detection model is employed to localize abrasive flaps. To overcome class imbalances, a hybrid approach is implemented where the neural network performs a robust one-class detection, while the final distinction between intact flaps and stubs is achieved through deterministic, scale-invariant post-processing filters. The system was validated on a dataset including various tool diameters, grit sizes, and wear conditions. For tools within the trained distribution, the system achieved 96.5% classification accuracy and a Mean Absolute Error of 0.14 and 0.12 for flap and stub counts. However, performance is limited when encountering significant geometric domain shifts, such as larger tool diameters. The results demonstrate that the combination of deep learning and geometric heuristics provides a robust basis for automated flap tear detection in industrial flap wheel grinding applications.

Keywords Abrasive flap wheel · Flap tear · Deep learning · Object detection

1 Introduction

Abrasive flap wheels are widely used for repair tasks or finishing of complex workpieces due to their high compliance [1, 2]. However, this compliance comes at the cost of high mechanical stress on the individual flaps, making tool wear a critical factor for process stability. In addition to grit wear and geometric change of the flap profile due to flap load, flaps may tear and detach prematurely. Ongoing investigations indicate that flap tears increase vibrations and localized loss of tool stiffness due to missing mutual support between flaps. While single flap tears have no significant effect on material removal and uniformity, this effect is intensified if adjacent flaps tear (tear cluster). Depending on the tool size, replacement is advisable once a critical number and spatial distribution of flap tears has occurred.

✉ Falko Kähler
f.kaehler@tuhh.de

1 Hamburg University of Technology, Institute of Aircraft Production Technology, Denickestr. 17, Hamburg 21073, Germany

To automate this decision-making process, an inspection system must be developed to detect and localize missing flaps or their remains on the flap wheel. Knowledge regarding the number and spatial distribution of these torn flaps, as well as their expected effects on the grinding result, then form the basis for deciding whether to continue using or replace the tool. However, the flap wheel structure and its properties pose challenges for classical image processing techniques due to stochastic variations in its visual appearance.

To overcome these limitations, this paper proposes a novel flap tear detection method utilizing a hybrid deep learning-based object detection approach. By combining a domain-specific image preprocessing with an Oriented Bounding Box (OBB) algorithm and a physics-informed filter logic, the proposed system robustly identifies and localizes flap tears.

The remainder of this paper is structured as follows: Section 2 introduces abrasive flap wheels and discusses the current state of research regarding tool wear and tool condition monitoring. In Section 3, the detection system for flap tears is conceptualized and the processing pipeline is introduced.

Published online: 15 August 2026

Section 4 describes development and implementation, including generation of training data and post processing of the detections. In Section 5, the detection pipeline is evaluated, while Sections 6 and 7 summarize the results and provide an outlook on future research.

2 State of the art

2.1 Flap wheel wear

Abrasive flap wheels, as depicted in Fig. 1, consist of radially mounted abrasive flaps on a core, fixed on a metal shaft. The flaps consist of a substrate (usually textile fabric) to which the abrasive grains are adhered using a binder. This structure corresponds to coated abrasives, giving the flap wheel its flexibility and shape adaptability [1, 3], but also unique wear mechanisms.

These wear patterns can be divided into microscopic wear of the abrasive grit as well as into macroscopic wear of the tool geometry. Grit wear has been investigated by Yang et al. [4], who identified three wear stages with different proportions of wear mechanisms (blunt, stepped fracture, cleavage fracture, pullout and adhesive wear). Regarding the geometric wear, Tönshoff et al. [5] investigated unmounted flap wheels during outer circumferential grinding and examined the flap length degradation, which leads to changing technological parameters, i.e. the effective cutting velocity and workpiece overlap decrease with decreasing diameter. In order to compensate the wear, an increased rotation speed is proposed. In contrast to even flap length degradation, uneven flap wear resulting from curved workpiece geometries affects the removal uniformity and could be compensated by adjusting the tool path [6]. Regarding flap tears,

Fig. 1 Abrasive flap wheels with torn flaps

Rahman et al. [7] investigated the mutual flap support using high speed imaging and noticed a sharp flap bend near the flap fixture during high engagement. The authors suspect that this cyclic bending and straightening can lead to premature flap failure. However, the precise causes, sequence and effects of flap failures were not studied.

2.2 Grinding tool wear monitoring and inspection

In grinding processes, traditional indirect tool condition monitoring heavily relies on process signals, such as cutting forces, acoustic emission, or spindle current, to infer process stability or tool wear [8–10]. While these indirect methods allow for continuous in-process monitoring, their signals are often ambiguous and highly susceptible to process noise, kinematics, and material variations. Consequently, indirect monitoring struggles to isolate exact geometric wear states or localized anomalies without a reliable, pre-established ground truth.

To overcome these inherent limitations, direct optical inspection has become increasingly critical. Unlike indirect signals that merely infer tool conditions, direct methods explicitly quantify the condition of the abrasive tool. This direct assessment provides an absolute reference that is essential for evaluating wear.

Driven by this need for objective wear quantification, automated direct inspection has been successfully established for a variety of conventional abrasive tools. For rigid grinding wheels, camera-based machine vision systems are frequently used to measure macroscopic 2D profile loss [11]. More advanced approaches utilize high-resolution imaging combined with machine learning to classify surface loading and localized grit wear directly on the wheel's topography. Similarly, automated optical inspection systems have been developed for abrasive belts to detect continuous grain wear and widespread degradation [12–14]. A shared characteristic of these conventional tools is a continuous, relatively predictable abrasive surface that allows for straightforward linear or rotational scanning.

When evaluating localized, discrete tool features, direct optical inspection has also seen significant advancements in the domain of geometrically defined cutting tools, such as milling cutters or turning inserts. For these tools, high-resolution machine vision combined with deep learning is widely utilized to classify or quantify localized wear phenomena like edge chipping or flank wear [15–17]. However, the success of these methods relies on the tool's rigid, deterministic geometry, which guarantees a predictable location of the cutting edge and an unobstructed line-of-sight for the camera.

Transferring established direct inspection concepts to compliant and stochastically structured tools such as flap wheels reveals a dichotomy in wear assessment. On a macroscopic level, evaluating the overall flap profile from a radial perspective to quantify global diameter reduction is methodologically comparable to the 2D contour inspection of rigid grinding wheels. However, assessing flap tears in radial imaging is severely hindered by the tool's overlapping arrangement of flaps, causing severe visual occlusion. Adopting an axial inspection perspective avoids flap occlusion, but introduces computer vision challenges due to stochastic visual tool appearance. Consequently, conventional contour inspection algorithms are rendered ineffective. This leaves a gap in the automated condition monitoring of compliant abrasives, necessitating a new inspection approach capable of resolving these structural peculiarities to quantify and localize individual missing flaps.

2.3 Conclusion

An examination of the current state of research reveals a gap in the understanding and automated inspection of flap wheel wear. While grit wear, uniform flap degradation, and selected flap profiles have been investigated, flap tears have only been briefly mentioned in the literature. The precise causes and effects of flap failures, as well as methodologies for their automated visual detection have not been addressed yet. Furthermore, while continuous macroscopic wear such as uniform flap shortening can be effectively mitigated by adjusting global process parameters, compensating the expected local tool stiffness reduction for discrete flap tears is fundamentally different. Due to the high rotational speeds typical in grinding applications, actively adjusting the tool engagement depth or force within a fraction of a single tool revolution and synchronous with flap tears is unfeasible. Therefore, tool replacement upon reaching a critical condition remains the only viable strategy to maintain process stability.

Consequently, the detection of flap tears must serve as a basis for condition-based tool replacement. While in-process monitoring via indirect signals is a common practice for conventional grinding processes, it is currently not feasible for flap tears. The fundamental reason is that the exact macroscopic effects of individual stochastic flap tears on the grinding process remain largely unresearched. Without a reliable ground truth to correlate a missing flap to a specific signal anomaly, indirect monitoring systems cannot be employed effectively. Therefore, a direct, intermittent optical inspection to evaluate the tool's physical condition between grinding cycles is necessary. While direct optical inspection is a well-established practice for geometrically deterministic cutting tools, the stochastic nature of abrasive flap wheels will likely demand a specialized approach. Since conventional image processing algorithms generally rely on rigid geometries and predictable contours, they are expected to struggle significantly with the flexible, overlapping, and visually chaotic structure of these compliant tools. Thus, to meet the demands of modern zero-defect manufacturing, exploring visual inspection driven by deep learning presents a highly promising alternative [18].

3 Conceptual design of the detection system

3.1 Image acquisition setup

Torn flaps are primarily identifiable visually by flap remnants ("stubs") left at the core, while secondarily, the distance between adjacent intact flaps tends to increase. In a radial view used to examine tool diameter and flap profile [6], the stubs and the gaps are usually occluded. However, the complete radial structure including flaps and stubs is recognizable from an axial viewing perspective. To acquire these images, the flap wheel is positioned axially in front of a stationary camera by an industrial robot. For a repeatable positioning of the flap wheel, the camera and endeffector are calibrated to the robot. However, the robot spindle rotation angle is not controlled and will be random. A horizontal ringlight aligned to the flap wheel is used to achieve a circumferential dark-field illumination. This configuration evenly illuminates the entire flap structure while preventing reflections from the flap wheel core. In addition, stray ambient light is effectivly supressed. The acquisition setup used in this work is depicted in Fig. 2 and is conceptualized as an intermittent inspection station located adjacent to the grinding cell. In an industrial deployment it is advised to physically separate the acquisition setup from the machining zone to avoid dust contamination, as well as to enclose it to also exclude any stray light. The robotic grinding setup in this work is physically restricted to tools up to 50 mm in diameter and the aqcuisition setup is designed for flap wheel sizes operational within the grinding cell. For tools $\geq$ 60 mm, some of the flap tips are already outside the field of view.

3.2 Geometric normalization and problem space reduction

Despite the axial inspection perspective, the radial arrangement and arbitrary angular position and orientation of the flap wheel structure pose a significant challenge for direct image analysis, as it complicates feature extraction and would necessitate complex, rotation-invariant algorithms. In the original circular image, the continuously changing spatial orientation of the target features precludes the efficient use of classical computer vision tools, such as

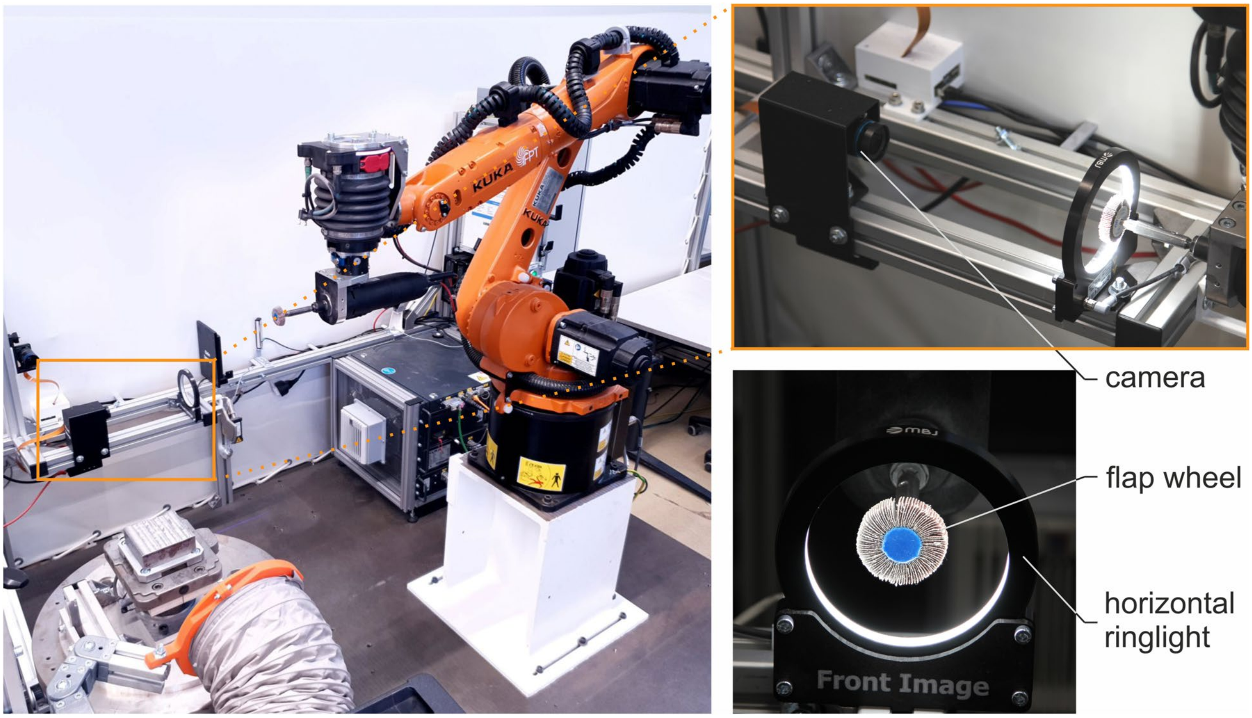

Fig. 2 Frontal image acquisition setup in a robotic grinding cell

directional gradient filters (e.g. edge detectors) or projection profiles. Similarly, Convolutional Neural Networks (CNNs) are inherently translation-invariant, but struggle with this rotational variance. Therefore, a geometric normalization to simplify the problem space is a fundamental prerequisite, regardless of whether classical image processing or deep learning techniques are subsequently applied.

The core of this normalization is a transformation of the initial frontal image into a polar coordinate system. Given a previously determined tool center point as the origin and by mapping the radial and angular information onto cartesian axes ($r \rightarrow y$, $\phi \rightarrow x$), the radial structure of the flap wheel is "unrolled" into a linear band structure. The goal of this transformation is to exploit the expected periodicity of the flaps, as the flaps ideally appear as vertical structures in the transformed image. In theory, this reduces the detection of anomalies (missing flaps) to an analysis of deviations within this periodic pattern.

In reality, however, the flaps are rarely attached with perfect spacing or aligned strictly radially to the core; they often appear locally slanted, curved, or twisted. To facilitate subsequent processing and support feature extraction further, the flaps should be aligned as vertically as possible using a shear transformation. While this transformation distorts the natural curvature of the flaps, such distortion is irrelevant for the detection of stubs. Since the initial inclination and curvature are highly specific to each individual tool, the optimal shear factor must be determined adaptively. In this work, the averaged horizontal image variance is computed across a predefined interval of shear factors. The maximum variance indicates the overall strongest vertical alignment of the flaps, thereby identifying the optimal shear factor for the respective tool.

By geometrically 'unrolling' the radial structure into a cartesian coordinate system, the complex rotational variance is universally transformed into a simpler translational variance. Combined with the shear transformation, this standardization forces all flaps into a uniform, vertical orientation, ensuring structural comparability and thereby drastically reducing the computational complexity and feature space for any downstream analytical or learning-based algorithm. The procedure for the geometric normalization of the frontal image is shown in Fig. 3.

3.3 Limitations of classical image processing approaches

Resulting from the shear transformation, the image exhibits a certain degree of horizontal flap periodicity. Classical analytical image processing approaches may now determine color or intensity profiles to calculate flap distances and deviations within the periodic pattern, thereby identifying gaps and torn flaps.

However, evaluating distances or periodicity of edge progressions, color and intensity values quickly proved insufficient, as the visual appearance of flap wheels is

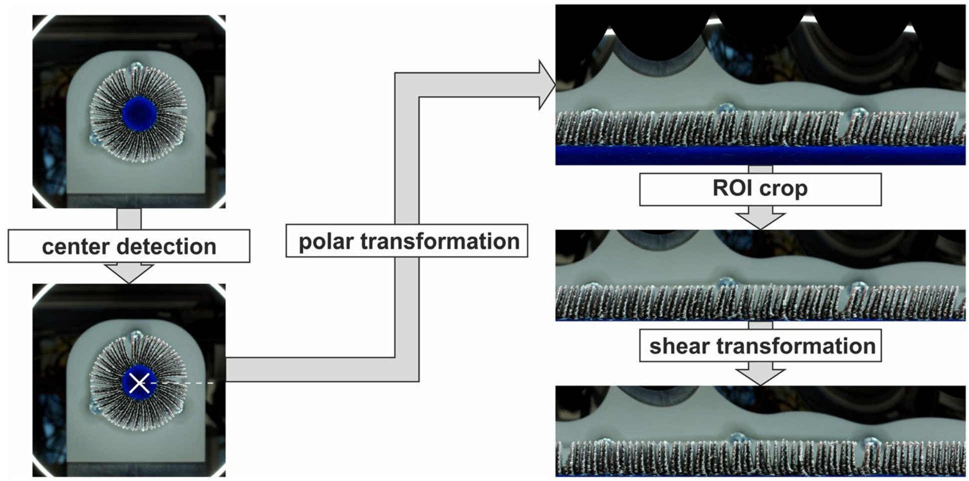

Fig. 3 Geometric normalization of the front image

subject to considerable fluctuations. The reasons for this include manufacturing tolerances, the flap structure and the resulting curvature due to flexibility, the varying number and dense packing of the flaps, occlusions by protruding fibers, as well as inhomogeneous coloring and irregular flap textures. This stochastic variance in appearance cannot be fully compensated for despite the described acquisition setup and image transformations. In addition, individual torn flaps do not necessarily result in a gap significantly larger than the natural variance between intact flaps. While larger gaps reliably indicate missing flaps, estimating the number of missing flaps based solely on distance measurements proves inaccurate.

For robust detection of missing flaps, it is therefore necessary to directly detect the distinct visual features of intact flaps and stubs. Secondary features (periodicity, distances, specific colors, etc.) and manually defined feature extractors (so-called "hand-crafted features") do not achieve the necessary robustness in this application. Zou et al. [19] describe this limitation in their historical review of object detection and highlight the paradigm shift brought about by Deep Learning, where Convolutional Neural Networks (CNNs) can directly learn invariant and highly complex feature representations from raw data to detect objects.

3.4 Learning-based flap detection

To recognize torn flaps and their location, a transition from pure image classification to instance-aware object detection is required. As Liu et al. [20] explain, generic object detection enables the precise localization of individual instances within an image. Since the geometric length measurement of each individual piece of fabric is essential for detecting torn stubs, the prediction of bounding boxes represents the most efficient bridge between visual recognition and subsequent analysis.

However, despite the preceding geometric normalization, the stochastic nature of the abrasive fabric causes individual flaps to retain local curvatures and slight individual inclinations. For this densely packed arrangement, the classical approach of axis-aligned bounding boxes (Horizontal Bounding Boxes, HBB) is insufficient. In such cases, a horizontal box inevitably encloses large parts of the background or neighboring instances, leading to spatial misalignment and severe noise during feature learning ("feature misalignment") [21]. For this reason, the use of Oriented Bounding Boxes (OBB) is highly advantageous, if not essential, for this problem area. OBBs better conform to the local physical rotation and curvature of the flaps, as well as the arbitrary bending of unsupported stubs, thereby allowing for a cleaner separation of the densely packed individual objects.

Among the available, pre-trained OBB detectors, the YOLO (You Only Look Once) architecture [22] represents a suitable choice for adaptation to this use case. YOLO belongs to the family of so-called one-stage detectors. Recent systematic reviews emphasize that the evolution of single-stage detectors, particularly the transition towards lightweight architectures like YOLOv8, has established them as the preferred state-of-the-art framework for real-time processing and edge-device deployment [23]. In contrast to more computationally intensive two-stage detectors (such as Faster R-CNN), which generate and filter region proposals in a separate, preceding step, YOLO predicts the bounding boxes and class affiliations in a single, efficient network pass [19, 20]. This compact architecture simultaneously processes the global image context and provides the real-time capability required for automated inline inspection at industrial production speeds (typically requiring less than 20 milliseconds per inference on standard industrial hardware), while maintaining high localization precision.

To further maximize the performance of the YOLO network, the problem is reduced to a pure 1-class detection. Since intact flaps and defective stubs physically consist of the same material and exhibit identical texture characteristics, semantic separation by the neural network would be prone to errors and highly dependent on absolute scaling. Furthermore, stubs occur rarely under normal operating conditions and are significantly outnumbered by intact flaps; a 2-class approach would thus result in a severe class imbalance, causing the rarer defect class to be learned poorly. Therefore, the model is trained exclusively to localize the presence of "abrasive fabric" using an OBB.

To finally differentiate between an intact flap and a flap stub, a hybrid architecture is implemented. Unlike conventional end-to-end deep learning classifiers that act as black boxes, this approach shifts the classification to a subsequent algorithmic step. This explicit separation directly addresses the critical industrial demand for transparent and traceable decision-making, as the deployment of pure black-box models in manufacturing is frequently hindered by their lack of interpretability [24, 25]. Since the preceding geometric normalization aligns the objects, the spatial extent of the bounding box directly correlates with the physical length of the flap. Consequently, a flap is classified as torn if and only if its localized geometry falls below a flexibly adjustable, dynamically calculated size threshold. This physics-informed approach not only solves the problem of subjective human labeling errors without the need to retrain the network, but it also guarantees that every defect classification provides a verifiable, explicitly measurable rationale. Ultimately, this ensures high fidelity, transparency, and comprehensible decision-making for industrial applications.

Regarding the specific choice of model generation, the decision was made to use YOLOv8 OBB instead of more recent versions. This decision is primarily based on the architectural control over object overlaps and the saturation of feature extraction. The evolutionary leaps of newer YOLO iterations aim to improve performance on highly complex, multi-class datasets [19]. For the single-class problem with constant texture presented here, YOLOv8 already achieves almost flawless feature extraction; parametrically heavier models are not expected to provide measurable information gain in this context.

Crucial to the choice of YOLOv8, however, is the handling of dense object clusters when processing image tiles. Because the high-resolution panorama images must inevitably be divided into overlapping sub-images (tiles), box duplicates are inevitably created during later fusion. Newer architectural approaches (like YOLOv10) abandon the Non-Maximum Suppression (NMS) algorithm to remove more uncertain duplicates in favor of dual-label assignment during training, which deprives the user of mathematical control over overlap resolution. YOLOv8 OBB retains NMS control as an explicit parameter in the architecture, allowing the Intersection-over-Union (IoU) threshold in the inference pipeline to be adapted to the physical overlap of the tile edges and the flaps. Complemented by the unmatched stability of the export routines for OBB angle calculations on industrial hardware, YOLOv8 provides a robust foundation.

3.5 Processing pipeline

The complete processing pipeline for the frontal images regarding missing flaps is shown in Fig. 4 and begins with geometric normalization. For the described polar transformation, it is first necessary to determine the tool center point. Subsequently, the frontal image is transformed and initially cropped to a fixed height that covers all considered tool dimensions. Afterwards, the image is sheared, using the best shear angle/factor determined for vertical flap alignment.

The geometric transformation of the radial tool structure into a linear panorama image inevitably creates artificial cut edges. Flaps or stubs that physically lie exactly on this dividing line are thus halved and distributed between the left and right edges of the image. The YOLO network would no longer be able to recognize these fragments as a coherent object, leading to detection gaps at the image edges. To compensate for this edge effect, the unrolled image is horizontally expanded using circular padding (wrap-around padding) immediately prior to inference. A defined pixel area from the left edge of the image is copied to the right edge and vice versa. This artificial overlap restores the continuous ring structure for the neural network, ensuring that every object on the tool circumference is depicted fully and uncut at least once.

However, due to the extreme aspect ratio of the unrolled panorama image, it cannot be passed as a whole to the YOLO network. Directly scaling it to the network's standard square input resolution would massively compress the image horizontally, resulting in the loss of the fine texture features and edge distances of the densely packed flaps. To preserve the native spatial resolution and morphological properties, the

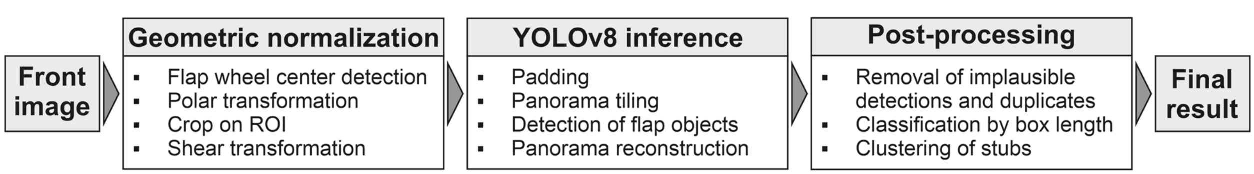

Fig. 4 Processing pipeline for flap tear detection

panorama images for training and inference are therefore divided into smaller, square tiles using a sliding-window approach. To prevent objects from being cut at the inner tile boundaries and thus overlooked, the tiling is performed with a defined overlap. This guarantees that every flap and every stub is fully visible in at least one tile and is captured by the network with native edge sharpness.

Following the YOLO inference and panorama reconstruction, the detections are post-processed. Duplicate detections, which inevitably arise from the padding and tiling of the panorama, are algorithmically fused back into a single object instance or removed in the downstream filter pipeline, considering each model's detection confidence. In addition, implausible detections are removed, and a distinction is made between an intact flap and a stub. Finally, the proximity of stubs is analyzed and directly adjacent stubs are merged into clusters.

4 Development and implementation

4.1 Construction of the training dataset

For the development of the Deep Learning-based flap detection, a training dataset was constructed that maps the specific geometric and optical properties of flap wheels. Panorama images of the frontal views preprocessed with the previously described geometric normalization serve as the data basis. These panorama images contain flap wheels in various diameter sizes (20-50 mm), widths (5-20 mm), grit sizes (grit 60-240), wear conditions (new, worn) as well as with or without missing flaps, and varying lighting conditions. In addition to tool images, individual polar-transformed background images were deliberately added to improve precision and reduce erroneous detections originating from the background.

The polar-transformed images were manually labeled using the Matlab Image Labeler, initially strictly differentiating the objects into two classes (intact flaps and flap stubs). Both classes exhibit the same flap thickness within an image but differ significantly in their length and visual appearance. During annotation, directly drawing the oriented bounding boxes manually was deliberately avoided. Instead, objects were marked with a polyline that follows their orientation and, if applicable, curvature. Subsequently, the lines are converted into the YOLO OBB format, which contains the object's class as well as the normalized image position and orientation of the bounding box for each mark. This procedure accelerates and simplifies the manual labeling process. Estimation errors regarding the position and orientation of the boxes are thus only caused by inaccurate nodes of the polylines.

The orientation of a bounding box is determined during the subsequent conversion process by a line of best fit through the marked nodes. The length of the box is then obtained by projecting the first and last node onto this line. The width of the bounding box can be either statically defined or dynamically determined during the conversion process. Since the flap width in the preprocessed image varies depending on the number of flaps, a constant box width could be too narrow or too wide for the actual flaps in some cases, causing information on flap edges and gaps to be lost or multiple flaps to be included. Therefore, the box width is calculated dynamically for each image based on the median horizontal distance between the center points of the nodes. The box width is subsequently set to 90% of this median distance so that the flaps are completely framed and parts of the gaps between the flaps are included. Minor overlaps of the boxes with closely adjacent flaps are accepted in this process. Finally in the conversion process, both labeled classes are combined into a single class. Figure 5 illustrates the labeling process. To visualize the stubs, they have been marked in red.

The final dataset comprised 74 panorama images. Since the YOLO model scales input images to a fixed square input size, the panorama images were each subsequently converted into square-shaped tiles (60% overlap) after labeling. This prevents the flaps from being extremely compressed horizontally down to sub-pixel size. Texture features and the aspect ratio of the objects are thus preserved. In total 888 individual tiles were generated (see Table 1).

4.2 Implementation and training

To determine suitable training parameters, model sizes, and filter values, various training configurations were first tested. A k-fold cross-validation with 5 folds was applied for this purpose. The dataset is randomly divided into k equally sized groups ("folds"). The model is retrained k times, with a different fold serving as validation data and the remaining folds as training data in each iteration. Finally, the results of all k runs are averaged to provide a robust overall value for model performance. This strategy offers several advantages over single training runs with a random data split. Because each data point is used once for testing and k-1 times for training, data utilization is maximized. Moreover, since the model is tested on different data combinations, the results are less susceptible to chance in the data split, thus reducing variance. Furthermore, the folds facilitate the detection of overfitting, as it is possible to assess whether the model only works well on a specific subset of data or actually generalizes. To prevent data leakage within the folds, the panorama images and all their corresponding tiles were strictly assigned to either the training or validation set. This strict

Fig. 5 Flap labeling procedure

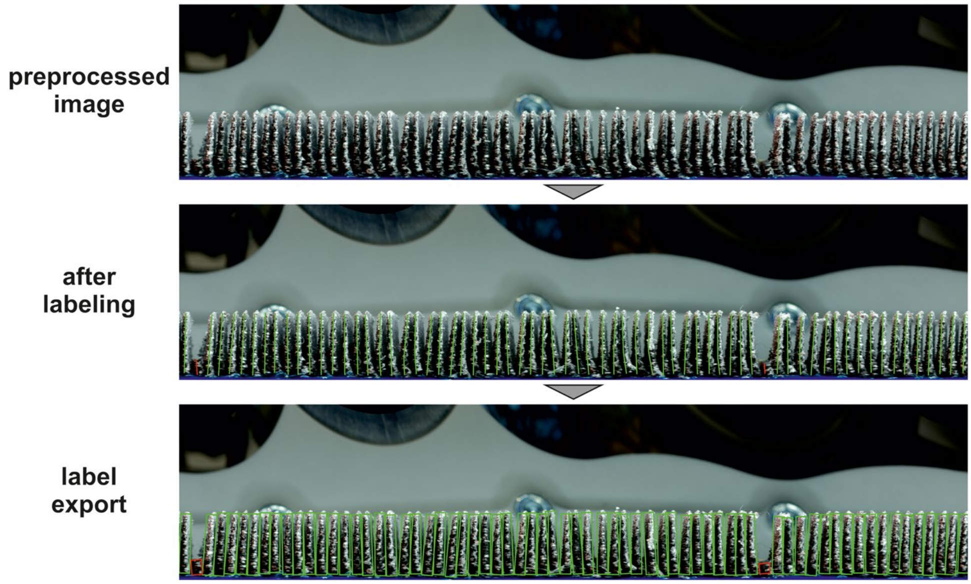

Table 1 Overview of the training dataset

Dataset Metric	Count
Total number of panorama images	74
↪ of which images with only intact flaps	21
↪ images with at least one defect (stub)	50
↪ images with background only	3
Tile overlap	60%
↪ tiles per panorama image	12
↪ total number of tiles	888
Total number of annotated objects	4749
↪ Original label: intact flap	4434
↪ Original label: stub	315
Ø Annotated objects per image	66.9
Max. number of stubs in a single image	21

separation ensures that the model does not artificially inflate its performance by memorizing shared background textures or specific tool features present in both sets. The entire software pipeline, including model training and evaluation, was implemented in Python utilizing the PyTorch deep learning framework. During this cross-validation phase, comprehensive hyperparameter tuning was conducted to adapt the baseline YOLOv8 architecture to the specific requirements of flap tear detection. It emerged that the "small" model size (YOLOv8s-OBB) with an input resolution of 640 pixels offers the optimal balance between detection performance and computational efficiency. Larger models did not achieve any significant performance advantages, but increased memory requirements and reduced inference speed. The deliberate choice of the YOLOv8s architecture restricts the overall model complexity (approx. 11 million parameters), effectively preventing the over-parameterization and overfitting often associated with deeper networks. This constrained complexity, combined with the comprehensive k-fold cross-validation strategy, significantly reduces model parameter uncertainty. Within the network's convolutional blocks, the SiLU (Swish-1) activation function was utilized to introduce the necessary non-linearity while effectively mitigating the vanishing gradient problem.

For the optimization of the network weights, the Stochastic Gradient Descent (SGD) optimizer was employed in combination with a cosine annealing learning rate scheduler (initial learning rate 0.001). This scheduler dynamically reduces the learning rate following a cosine curve, facilitating a smooth convergence into the global loss minimum towards the end of the training process. The overall optimization objective was defined by a combined loss function comprising Binary Cross-Entropy (BCE) for class prediction, Distribution Focal Loss (DFL), and an Intersection over Union (IoU) based regression loss for the oriented bounding boxes. Given the single-class approach where precise geometric localization of the flap boundaries is significantly more critical than semantic classification, the bounding box regression loss was heavily weighted in the hyperparameter configuration compared to the classification loss.

To achieve high robustness and generalization capability of the model towards unknown tools, massive data augmentation—such as scaling, rotation, mosaic generation, and color jitter—was incorporated into the training process. This forces the model to learn abstract features instead of concrete shapes, positions, orientations, or colors of the flaps. In addition to data augmentation, structural regularization techniques were implemented to actively combat overfitting. These included Weight Decay (L2 regularization) to penalize excessively large network weights, Label Smoothing to reduce model overconfidence, and Dropout layers, which randomly deactivate a fraction of neurons (10%) during training to enforce redundant and robust feature representations. Furthermore, early stopping was utilized during the initial k-fold validation to monitor the validation loss and identify the optimal training duration before overfitting occurs. Based on these insights, the final model was trained for a fixed, saturated duration of 250 epochs.

After the most suitable model architecture, hyperparameter set, and training duration were determined, the final model was trained. For this, 100% of the training data was used, providing the network with the largest possible training base. The validation of the model and the pre- and post-processing take place on a separate validation dataset that is not used in the training process.

4.3 Post-processing of detections

The downstream filtering algorithms subsequently evaluate the bounding boxes output by the object detection. To clear redundant or physically implausible bounding boxes, a multi-stage filtering logic was implemented, whose thresholds were, wherever possible, derived data-driven from the annotated training dataset. In addition, stubs that lie next to each other are clustered.

Stage 1

In the first stage, a plausibility check of the boxes is performed. Here, the boxes must

- exhibit a minimum width based on the median box width,
- have a longitudinal axis at an angle of less than 45° from the vertical, and
- begin near the root area of the flaps.

During development, the minimum width was set to 50% of the median box width, with the median box width calculated from all detected boxes. The deviation of the longitudinal axis of the boxes is restricted to 45°. This value was determined based on the training data and the maximum box inclination observed there (41.3°), thus empirically filtering out illogical boxes. The final condition filters out boxes that do not start near the core. This removes floating boxes that do not mark the physical connection of the flaps to the core.

Only boxes that meet all conditions are considered plausible. These are passed to the subsequent stages and used to calculate the median box length in Stage 3.

Stage 2

In Stage 2, duplicate detections resulting from padding and overlapping tiles from the panorama or YOLO predictions are removed. A box is removed,

- if it overlaps too much with another box and the other box has a higher confidence,
- if its center is too close to another box and that box has a higher confidence, or
- if the box is heavily enclosed by another box.

The evaluation of the training data revealed a maximum natural overlap (Intersection over Union, IoU) of adjacent flaps of 0.16. The NMS threshold for the maximum permissible overlap was consequently set to 0.20. Additionally, the data analysis showed a minimum natural distance between box centers of 17.6 pixels, from which a minimum distance of 13 pixels was derived for post-processing to reliably discard duplicates that are too close together. To resolve duplicate detections at the padding seams, the X-coordinates (circular angle) of the bounding boxes are projected back into the original image space via a modulo operation before filtering (circular distance calculation). Through this virtual overlay of the edge areas, seam duplicates can be seamlessly processed and eliminated by the pipeline's regular NMS.

While the classic IoU filter works excellently for boxes of similar scaling, it reaches mathematical limits with highly asymmetrical object sizes. Since the union of both boxes is in the denominator for the IoU calculation, the overlap value is artificially low for a very small and a very large box, even if the small box is completely enclosed by the large one. Consequently, a small erroneous box completely nested inside a large box yields an IoU significantly below the NMS threshold, preventing its deletion.

To eliminate such erroneously detected artifacts (e.g., alleged stubs detected within the texture of intact flaps), an additional Matryoshka filter (Intersection over Smaller Area) was added to the pipeline. This evaluates the overlap not based on the union, but exclusively relative to the area of the smaller object. While thresholds for the minimum distance or the classic IoU can be derived directly from the physical geometries of the manually annotated training data (ground truth), the Matryoshka factor requires an empirical approach. The phenomenon of nested boxes does not occur in reality and is only an inference artifact of the neural network (e.g., due to misinterpretation of local texture deviations on intact flaps). Since such inclusion errors do not exist in the error-free ground truth, the threshold was empirically set to 33% based on physical plausibility. This value provides sufficient tolerance for the optical overlap of densely packed flap edges but reliably eliminates illogical, network-generated ghost detections within the flap areas before the final classification stage.

Stage 3

In the third stage, a final distinction is made between an intact flap and a stub for each box. The essential physical distinguishing feature between flaps and stubs is their difference in length. Therefore, an object is classified as a flap/stub,

- if the length of the box is above/below a threshold value.

Since the length ratio between intact flaps and stubs depends on the tool design (diameter and flap length), a threshold value that is as universally applicable as possible for all tool designs and wear conditions is of interest. To determine this threshold, the training dataset was evaluated with regard to the median lengths of intact flaps and stubs. This showed that stubs exhibit between 12 % and 55 % of the median intact flap length. For classification, a threshold of 57 % of the intact flap median box length is therefore selected. During inspection this median box length is calculated from the top 20% longest boxes of an image to prevent the calculation from being skewed if there are a lot of torn flaps. The classification threshold was deliberately set at 57 %, establishing a tight decision boundary just above the maximum observed stub length ratio. Choosing a higher threshold (e.g., 80 %) was avoided to cover uneven flap lengths which can occur through combined wear mechanisms of intact adjacent flaps. A highly restrictive threshold of 57 % maximizes the tolerance for this natural, uneven wear and accommodates slight inaccuracies in bounding box predictions, thereby effectively minimizing the occurrence of false-positive defect classifications.

4.4 Determination of flap tear clusters

The impact of spatially clustered stubs on the grinding result is the topic of ongoing research. First results indicate that clustered stubs, depending on the cluster size, are expected to be critical for tool behavior and have a negative impact on grinding results. Therefore, in addition to stub classification, stub clusters need to be determined for the decision whether to continue with the tool or replace it. Therefore, starting from classified stubs, neighboring stubs are searched. If another stub is found that is not separated from the current stub by an intact flap, these stubs are assigned to a cluster.

Figure 6 shows the final result of the implemented flap and stub detection in cartesian coordinates along with the position and, if applicable, the grouping of missing flaps.

5 Validation

To verify and validate the performance and robustness of the developed system for detecting flap tears, the agreement between the calculated metrics and the actual condition of the tool (ground truth) was evaluated. A validation dataset consisting of N=70 images of an equal number of individual flap wheels serves as data basis. The dataset was compiled to represent a high degree of variance: It contains both new and worn tools with varying numbers of torn flaps, ranging from single ones to massive clusters.

Furthermore, the dimensions (diameter, flap width) and grit size were varied. Here, designs with nominal diameters of 60 and 80 mm were also tested, which physically exceed the grinding hardware limitations and lie outside the learned distribution. These larger dimensions were deliberately withheld from the training phase to evaluate the system's robustness against severe geometric domain shifts (out-of-distribution testing). In fact, since a performance shift is expected here, the performance metrics are determined both for the entire validation dataset (N=70), the in-distribution subset (N=59) containing the learned distribution (up to 50 mm diameter) and out-of-distribution subset (N=11).

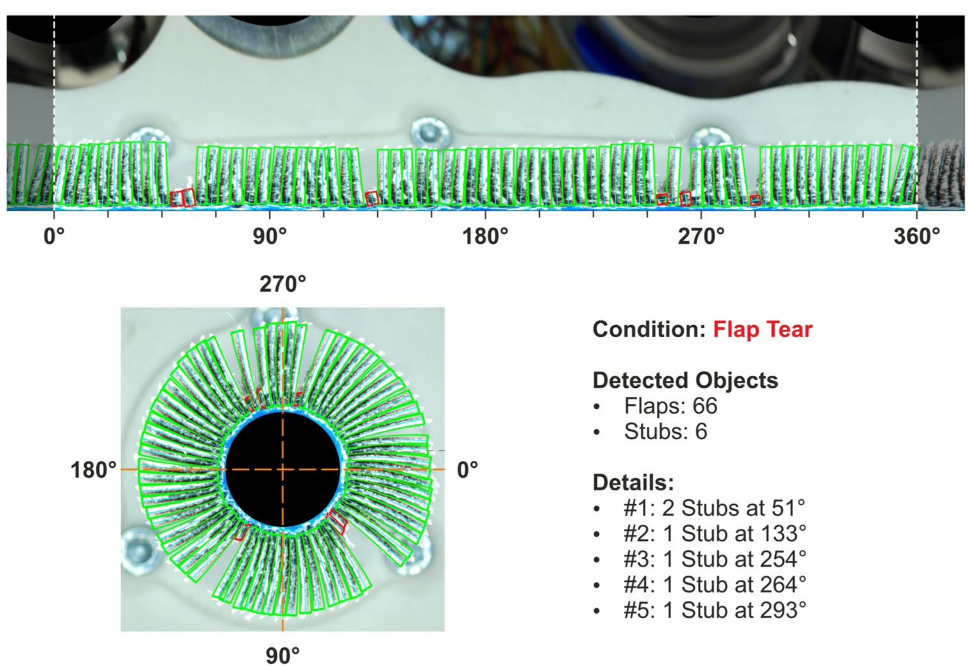

Fig. 6 Final detection results

For each image, a manual reference annotation (ground truth) was created, in which the number of intact flaps, missing flaps (flap stubs) and defect clusters were visually determined and documented. To investigate the generalization capability of the detection and avoid data leakage of specific tool instances, the validation dataset contains exclusively images of physical flap wheels that were not already part of the training. This object level separation ensures that the model recognizes abstract flap features rather than merely the specific texture or geometry of previously known tools.

5.1 Classification performance

First, the system's ability to distinguish between tools with exclusively intact flaps and tools with stubs was evaluated. The results of this classification are presented in Table 2 for the entire dataset as well as the "In-Distribution" and "Out-of-Distribution" subsets.

The confusion matrices not only prove the performance of the neural feature extraction but specifically confirm the robustness of the downstream classification logic. While the system achieved good results for the in-distribution dataset, misclassifications occurred especially with unknown tool sizes not part of the training dataset. As a result, the performance on the entire dataset decreased.

A detailed analysis of the misclassifications at unknown tool dimensions (out-of-distribution, 60 mm and 80 mm), as illustrated in Fig. 7, confirms the impact of a geometric domain shift. Because larger tool circumferences are projected onto the same tile sizes, the flaps experience significant horizontal compression in the pixel space. This massively alters the aspect ratio as well as the spatial texture frequency compared to the training data. Consequently, two primary failure modes occured: First, in particular for 80 mm samples, the object detection network partially spans bounding boxes only over fragments of intact flaps (e.g. just at the flap root), while implausible boxes are filtered (see Fig. 8). In the subsequent deterministic filtering logic, these incomplete boxes affect the length threshold calculation more than some cut-off flap tips, provoking false-positive tear classifications as well as missed stubs (false negative). Secondly, feature extraction tends to fail for severely compressed flaps and stubs, causing them to be overlooked by the network and thus omitted.

Table 2 Comparison of the classification results

Classification Metric	In-Distribution	Out-of-Distribution		Total Dataset
	∅ 20-50 mm	∅ 60 mm	∅ 80 mm	
No. images N	59	5	6	70
True Positives (TP)	30	1	1	32
True Negatives (TN)	27	2	2	31
False Positives (FP)	2	1	2	5
False Negatives (FN)	0	1	1	2
Overall Accuracy	96.61 %	60.00 %	50.00 %	90.00 %
Balanced Accuracy	96.55 %	58.33 %	50.00 %	90.11 %
Recall (Flap Tear)	100.00 %	50.00 %	50.00 %	94.12 %
False Positive Rate	6.89 %	33.33 %	50.00 %	13.89 %

5.2 Object quantification performance

A central quality characteristic is the correct number of flaps and stubs (total and adjacent). The number of missing flaps is particularly critical as these values serve as thresholds for tool replacement. To avoid statistical dilution through over- and undercounting, the counting accuracy was primarily evaluated using the Mean Absolute Error (MAE).

As listed in Table 3, the achieved MAE of 0.14 for intact flaps and 0.12 for flap stubs represents a remarkably high counting precision. Considering the high density of the tool structure, which features an average of 66.9 individual objects per flap wheel, an MAE of 0.14 corresponds to a relative counting error of approximately 0.2%. This low error rate indicates that the system operates almost error-free on the majority of tools, with miscounts limited to a maximum of a single object in rare edge cases of the in-distribution subset.

However, evaluating the quantification metrics across the out-of-distribution subsets highlights a severe performance degradation. While the model already suffers a significant loss of accuracy for the 60 mm samples — achieving an error-free count in only 60.0 % of cases, which falls short of the reliability required for practical application — the quantification capability collapses completely for the 80 mm diameter group (0.0% error-free counts). This degradation is further reflected in the drastic increase of the MAE from an already elevated 1.2 at 60 mm to 10.5 for intact flaps, 5.0 for flap stubs, and 3.67 for defect clusters at 80 mm. Rather than exhibiting a slight drop in precision, the model reaches its operational limits, failing partially at 60 mm and systemically at 80 mm.

Closer analysis of results with erroneous counts reveals different causes depending on the considered subset of the dataset. In the in-distribution dataset, no object was overlooked or filtered out in any case. Facilitated by an optical overlay of a stub and an intact flap as well as heavily frayed stubs with long fabric remnants, the counting errors resulted solely from incorrect class assignments.

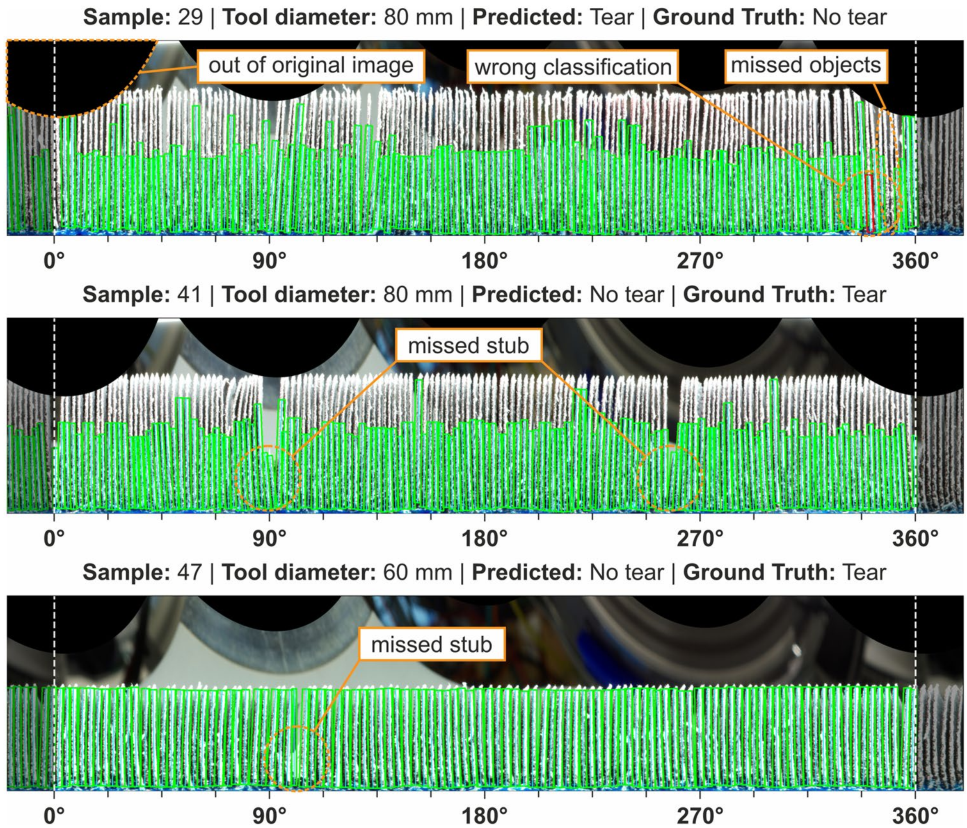

Fig. 7 False classification analysis of the out-of-distribution subset

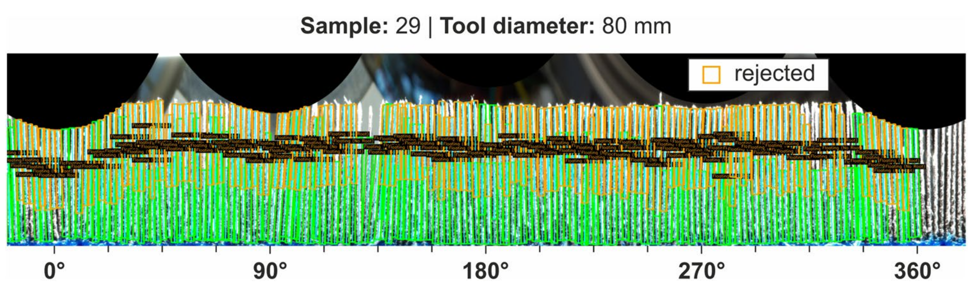

Fig. 8 Erroneous box filtering for out-of-distribution sample

Table 3 Comparison of the quantification results

Metric	In-Distribution	Out-of-Distribution		Total Dataset
	∅ 20-50 mm	∅ 60 mm	∅ 80 mm	
No. images N	59	5	6	70
Images with error-free count	48 (81.4 %)	3 (60.0 %)	0 (0.0 %)	51 (72.8 %)
MAE intact flaps	0.14	1.2	10.5	1.01
MAE flap stubs	0.12	0.6	5.0	0.53
MAE defect clusters	0.10	0.4	3.67	0.40

For the out-of-distribution dataset, many object boxes were rejected after YOLOv8 inference due to suspected implausibilities, as depicted for sample no. 29 in Fig. 8. This issue is particularly pronounced for the 80 mm samples, where the geometric deviation from the training distribution reaches its extreme and causes the elevated MAE values observed in Table 3. Due to the increasingly unfamiliar ratio of flap/box length to width to the object detection network, the boxes often cover only half the flap length. Particularly for 80 mm flap wheels, these bounding boxes rarely encompass the entire flap and either start at the bottom at the flap root or at the top at the flap tip, where these floating tip boxes are rejected. This impairs the length threshold essential for subsequent class assignment through too-short box lengths. Furthermore, coarser box overlaps occur due to the unusually thin flap structure of the larger diameters, leading to a substantially higher rate of discarded boxes for 80 mm samples and leaving several flaps entirely unassigned.

6 Discussion

During the validation of the developed missing flap detection system, the detection quality demonstrated a dependency on the tool dimensions represented in the training. If flap wheels are fed into the system whose diameter or flap length deviate significantly from the training data, a significant drop in model performance is observed. This effect can be attributed to a phenomenon known in pattern recognition as domain shift. The underlying neural network (YOLO) learned specific visual features – such as the spatial frequency of the texture, edge progressions, and lighting gradients – based on a defined distribution. If the geometric proportions in the image change significantly, the extracted image features fall outside this distribution. This leads to reduced confidence or wrong positioning of the initial bounding boxes. For implementing the detection approach, it is therefore recommended to cover the expected tool dimensions within the training dataset.

When evaluating the overall system, however, it is essential to differentiate between feature-based object detection and downstream deterministic evaluation logic. The filtering and classification pipeline developed in this work was deliberately designed to be scale-invariant and data-driven. The parameters for distinguishing between an intact flap and a defect are not based on absolute pixel values but on relative statistical size determinations, such as using the median of the longest detected objects as a dynamic reference length. Consequently, the evaluation logic scales automatically with the respective tool size.

However, the constant length threshold method used for classification may become increasingly susceptible to errors as overall tool wear progresses. Because the intact flaps shorten uniformly over the tool's operating life, their length gradually converges with that of previously formed flap stubs. Consequently, the relative threshold drops, which may allow longer stubs to exceed the limit and be falsely classified as intact flaps (false negatives). A potential solution to this limitation is the implementation of an adaptive threshold that increases over the tool's usage time, coupled to the current state of uniform flap wear measured against an initial baseline. From a process perspective, however, this geometric convergence also mitigates the severity of such errors. As the length differential between intact flaps and stubs minimizes, the physical and mechanical difference between them disappears. Consequently, misclassifications that occur due to this reduced separation are expected to have a smaller impact on the overall grinding result.

To enable the detection system for a broader range of tool sizes in the future, no structural adjustment of the evaluation architecture is required. Instead, performance can be increased through transfer learning methods (fine-tuning). By adding a relatively small amount of annotated image data of new tool dimensions to the existing training dataset, the detection model could be expanded iteratively, thereby reducing classification and counting errors. The physically based filter logic as the universal evaluation core can remain unchanged.

Furthermore, the training methodology of the feature extraction model warrants critical consideration. The current implementation relies on an SGD optimizer with a cosine annealing scheduler, which provided a robust baseline and high accuracy. However, this approach leaves potential performance gains untapped and contributes to model parameter uncertainty, as the influence of adaptive optimization algorithms (e.g., Adam, AdamW, or RMSprop) on this specific industrial dataset was not investigated. Since this study primarily focused on proving the feasibility of the hybrid geometric pipeline, fundamental neural network optimization was constrained. Addressing this limitation through a systematic benchmark and automated hyperparameter tuning remains a necessary step to maximize efficiency and minimize parameter uncertainty in future deployments.

Finally, a critical aspect of the system's practical deployment is its processing speed. Although the underlying YOLOv8 OBB model achieves its characteristic high-speed inference on a per-tile basis, the overall performance depends on the entire end-to-end processing chain. This includes image preprocessing, multi-tile inference per image, and the deterministic post-processing rules. Tests have shown that the system currently achieves a total processing time of approximately 1.5 seconds per tool on desktop hardware, without any specific optimizations such as TensorRT conversion or batch inference in the software implementation. While this unoptimized latency may be already sufficient for intermittent inspection tasks, transitioning to hardware-optimized execution environments and general software optimization could significantly reduce this timeframe.

7 Conclusion and outlook

In this work, a visual inspection of abrasive flap wheels in order to detect flap tears is proposed. The approach utilizes axial images of flap wheels, where flaps and stubs of failed flaps are best visually recognizable. These images are geometrically normalized, before a YOLOv8 OBB detection model is applied to detect flaps objects. The detections are subsequently processed to identify duplicates as well as plausible objects, and finally distinguish between intact flaps and flap stubs.

Results showed good performance both in terms of object classification and quantification. However, the results depended on the tool properties used for training.

When applied on unknown flap wheel sizes, flaps are often not fully recognized and bounding boxes not properly positioned, and thus partially filtered as implausible.

In the future, the detection approach could be linked to continuous process monitoring, in order to detect flap tear events and, if necessary, trigger the visual flap wheel inspection and decision on tool change. However, the decision must be based on an understanding of the processes involved in flap tears. To achieve this, correlating the automatically quantified flap tear condition with actual machining performance metrics, such as material removal rate, removal uniformity or surface roughness, will be the focus of future experimental studies.

Acknowledgements The research leading to these results received funding from the Federal Ministry for Economic Affairs and Energy (BMWE) of Germany in the project LEGATO under the Grant Number 20Y2205C as part of the Federal Aeronautical Research Programme LuFo VI-3. The authors are responsible for the content of this publication.

Supported by:

Federal Ministry for Economic Affairs and Energy

on the basis of a decision by the German Bundestag

Author Contributions Falko Kähler: Conceptualization, Methodology, Data Collection, Evaluation, Writing. Jessica Ehrbar: Review. Thorsten Schüppstuhl: Funding acquisition.

Funding Open Access funding enabled and organized by Projekt DEAL.

Availability of Data and Materials The data used for this research may be requested from the corresponding author.

Declarations

Conflicts of Interest The authors declare that they have no known competing interests (financial or personal) that are relevant to this article.

References

1. Zhang J, Shi Y, Lin X, Li Z (2017) Five-axis abrasive belt flap wheel polishing method for leading and trailing edges of aero-engine blade. Int J Adv Manuf Technol 93(9–12):3383–3393. https://doi.org/10.1007/s00170-017-0717-z
2. Huai W, Lin X (2020) An optimization study of polishing efficiency of blisk and its technological parameters. Sci Prog 103(3):36850420957850. https://doi.org/10.1177/0036850420957850
3. Zhu W, Beaucamp A (2020) Compliant grinding and polishing: A review. https://doi.org/10.1016/j.ijmachtools.2020.103634
4. Yang R, Lin X (2024) Wear behaviour of abrasive tools and its effect on surface integrity during polishing superalloy GH4169G with abrasive cloth flap wheels. J Manuf Process 120:286–304. https://doi.org/10.1016/j.jmapro.2024.04.053
5. Tönshoff HK, Hinzmann G (1985) Untersuchungen zum Lamellenradschleifen und Bandschleifen an Kontaktscheiben. Institut für Fertigungstechnik und Spanende Werkzeugmaschinen,
6. Kähler F, Ehrbar J, Schüppstuhl T (2025) Concave geometric wear compensation in automated flap wheel grinding. Lecture Notes Mech Eng 614–622. https://doi.org/10.1007/978-3-032-05610-8_61
7. Rahman MDH, Komath S, Subbiah S (2025) High-speed camera imaging-based investigations of compliance behaviour and flap interactions in coated abrasive flap wheels. Manuf Lett 44:498–506. https://doi.org/10.1016/j.mfglet.2025.06.059
8. Pandiyan V, Shevchik S, Wasmer K, Castagne S, Tjahjowidodo T (2020) Modelling and monitoring of abrasive finishing processes using artificial intelligence techniques: a review. J Manuf Process 57:114–135. https://doi.org/10.1016/j.jmapro.2020.06.013
9. Pimenov DY, Da Ribeiro Silva LR, Kuntoğlu M, Abrão BS, Dos Santos Paes LE, Linul E (2025) Review of advanced sensor system applications in grinding operations. J Adv Res 77:371–405. https://doi.org/10.1016/j.jare.2025.01.013
10. Caesarendra W, Triwiyanto T, Pandiyan V, Glowacz A, Permana SDH, Tjahjowidodo T (2021) A CNN prediction method for belt grinding tool wear in a polishing process utilizing 3-axes force and vibration data. Electronics 10(12):1429. https://doi.org/10.3390/electronics10121429
11. Su JC, Tarng YS (2006) Measuring wear of the grinding wheel using machine vision. Int J Adv Manuf Technol 31(1–2):50–60. https://doi.org/10.1007/s00170-005-0172-0
12. Oo H, Wang W, Liu Z (2020) Tool wear monitoring system in belt grinding based on image-processing techniques. Int J Adv Manuf Technol 111(7–8):2215–2229. https://doi.org/10.1007/s00170-020-06254-1
13. Wan Q, Zou L, Han C, Wang W, Qian K, Ou J (2022) A U-net-based intelligent approach for belt morphology quantification and wear monitoring. J Mater Process Technol 306:117652. https://doi.org/10.1016/j.jmatprotec.2022.117652
14. Ren L, Yan W, Wang N, Pang W, Zhang G (2025) A vision-based approach for precise wear evaluation of abrasive belts with irregular morphology in flexible grinding. Coatings 15(11):1257. https://doi.org/10.3390/coatings15111257
15. Bergs T, Holst C, Gupta P, Augspurger T (2020) Digital image processing with deep learning for automated cutting tool wear detection. Procedia Manufacturing 48:947–958. https://doi.org/10.1016/j.promfg.2020.05.134
16. Friedrich M, Gerber T, Dumler J, Döpper F (2023) A system for automated tool wear monitoring and classification using computer vision. Procedia CIRP 118:425–430. https://doi.org/10.1016/j.procir.2023.06.073
17. Schlager E, Windisch A, Hanna L, Klünsner T, Hagendorfer EJ, Feil T (2025) Evaluation of data augmentation and loss functions in semantic image segmentation for drilling tool wear detection. J Intell Manuf 36(2):1491–1503. https://doi.org/10.1007/s10845-023-02313-y
18. Dewi C, Pradibta Fury EK, Rosario Putra JS, Feodora Aritonang EJ, Setia Permana AB, Md Fauadi MHF, Aprius Sutresno S, Riantama D (2026) YOLO-based Visual Inspection in Industry

4.0: A Comprehensive Review of Smart Manufacturing Applications. Journal of Advanced Manufacturing Technology (JAMT) 20(1):81–114. https://jamt.utem.edu.my/jamt/article/view/7009

19. Zou Z, Chen K, Shi Z, Guo Y, Ye J (2023) Object detection in 20 years: a survey. Proc IEEE 111(3):257–276. https://doi.org/10.1109/JPROC.2023.3238524
20. Liu L, Ouyang W, Wang X, Fieguth P, Chen J, Liu X, Pietikäinen M (2020) Deep learning for generic object detection: a survey. Int J Comput Vision 128(2):261–318. https://doi.org/10.1007/s11263-019-01247-4
21. Wang K, Wang Z, Li Z, Su A, Teng X, Pan E, Liu M, Yu Q (2025) Oriented object detection in optical remote sensing images using deep learning: a survey. Artif Intell Rev 58(11). https://doi.org/10.1007/s10462-025-11256-0
22. Jocher G, Chaurasia A, Qiu J (2023) Ultralytics YOLOv8 (Version 8.0.0). https://github.com/ultralytics/ultralytics
23. Guleria A, Varshney K, Jindal S (2025) A systematic review: object detection. AI & SOCIETY 40:6511–6528. https://doi.org/10.1007/s00146-025-02372-0
24. Arinez J, Chang Q, Gao RX, Xu C, Zhang J (2020) Artificial intelligence in advanced manufacturing: Current status and future outlook. J Manuf Sci Eng 142(11):110804. https://doi.org/10.1115/1.4047855
25. von Rueden L, Mayer S, Beckh K, Bogdan M, Garcke J et al (2021) Informed machine learning - a taxonomy and survey of integrating prior knowledge into learning systems. IEEE Trans Knowl Data Eng 35(1):614–633. https://doi.org/10.1109/TKDE.2021.3079836