Forschung im Ingenieurwesen (2025)89:70
https://doi.org/10.1007/s10010-025-00833-y

ORIGINALARBEITEN/ORIGINALS

# A vision-assisted condition monitoring framework for wear monitoring of gearboxes

Djordy Van Maele[1,2] · Jean Carlos Poletto[1,2,4] · Roeland De Geest[3] · Wenzhi Liao[3] · Ney Francisco Ferreira[4] · Patric Daniel Neis[4] · Dieter Fauconnier[1,2] · Patrick De Baets[1,2]

Received: 15 September 2024 / Accepted: 4 April 2025


**Abstract**

This paper presents the development of a prototype system for in-situ wear monitoring of helical gears using vision-based condition monitoring (VCM). The research addresses the challenge of monitoring gearbox wear, particularly pitting, an inherent failure mechanism in high-power applications such as wind turbines. A VCM system was integrated into an FZG gear test rig to automatically capture images of gear surfaces during operation, using an oil splash mitigation strategy to ensure high image quality. These images were processed using an image processing pipeline, incorporating a deep learning Feature Pyramid Network for automated annotation of wear defects, trained based on an iterative approach to reduce annotation effort during the training phase. Hereafter, this paper explores metrics related to pitting evolution, including pit area, height, width, and location on the gear surface, as well as insights these metrics give on pitting evolution. This framework's ability to track the spatiotemporal development of gear wear provides valuable insights for condition monitoring and predictive maintenance in industrial applications

✉ Djordy Van Maele
djordy.vanmaele@ugent.be

Jean Carlos Poletto
jeancarlos.poletto@ugent.be

Roeland De Geest
roeland.degeest@flandersmake.be

Wenzhi Liao
Wenzhi.Liao@flandersmake.be

Ney Francisco Ferreira
neyferr@gmail.com

Patric Daniel Neis
patric.neis79@gmail.com

Dieter Fauconnier
dieter.fauconnier@ugent.be

Patrick De Baets
patrick.debaets@ugent.be

1 Soete Laboratory, Ghent University, Technologiepark Zwijnaarde 46, 9052 Zwijnaarde, Oost-vlaanderen, Belgium

2 Core lab MIRO, Flanders Make @ UGent, 9000 Ghent, Oost-vlaanderen, Belgium

3 Flanders Make, Oude Diesterbaan, 3920 Lommel, Limburg, Belgium

4 Laboratory of Tribology, Federal University of Rio Grande do Sul, Osvaldo Aranha 99, Porto Alegre, Rio Grande do Sul, Brazil

## 1 Introduction

Machines such as wind turbines require increased torque density, causing components such as their gearboxes to be stretched to their limits. Consequently, wind turbine gearboxes are increasingly more liable to premature failure due to need for high torque density. For each wind turbine gearbox that fails, this results in long downtime and high maintenance costs, which is especially problematic in industries such as offshore wind [1, 2].

Indeed, maintenance of this equipment cannot be performed at any given time due to extreme weather conditions at sea, and even when possible, it comes with challenges. Walford [3] estimated that replacing a gearbox in a 660 kW wind turbine costs up to €120,000, while for a 1.5 MW wind turbine, this value can be 3 to 4 times as much. Considering that modern offshore wind turbines are between 7–10 MW [4], the repair costs become significantly high. Therefore, industries like offshore wind highly benefit from accurate and robust condition monitoring. This allows for better insight into the health status of these gearboxes and the implementation of a predictive maintenance approach. This way, unexpected failure can be avoided. Secondly, remotely monitoring the health could allow wind turbine owners to perform temporary remote maintenance

Published online: 20 May 2025

actions when the wind turbine is inaccessible to extend its lifetime until it becomes accessible again for maintenance.

Currently, multiple techniques are used to monitor the condition of gearboxes. These techniques include oil monitoring [5, 6], vibration monitoring [7, 8], acoustic emissions [9, 10], microphones [11], motor current [12], temperature, load and speed sensors. All these techniques have in common that they attempt to estimate the damage by observing patterns and anomalies in the measured signals, which is an indirect approach. As a result, for applications such as wind turbines, once one of these techniques triggers a certain alarm threshold, a manual visual inspection is performed, which comes with high associated costs. This inspection is necessary to evaluate the gearbox's health and confirm the existence of the failure indicated by the indirect technique. If an integrated camera could monitor the gears instead, the turbine controllers could gain immediate insight into the current health status of their equipment without having to send out an operator for manual inspection. Therefore, it is necessary to investigate the use of vision-based condition monitoring techniques for in-situ and direct monitoring of the gear surface.

Gears are subjected to many gear failure modes during operation. However, not all gears are subjected to each failure mode. The ISO 10825:2021 standard [13] describes more than 40 different terms of gear failure modes. Among these, some are more relevant from a tribological point of view, such as abrasive wear, adhesive wear and fatigue wear (pitting). Adhesive wear is typically associated with high contact temperatures leading to lubricant film breakdown [14–16], and vice versa, ultimately leading to direct contact of the matching gear surfaces. Besides inadequate lubrication, during the start-up of machines, the gears could operate under boundary or mixed lubrication [17]. For example, low lubricant temperature can cause short periods of oil starvation, also leading to adhesive failure. Meanwhile, abrasive wear is commonly caused by contamination in the oil, such as wear debris, including those originating from other wear mechanisms. However, abrasive and adhesive wear can be mitigated using an adequate lubrication system and only tend to occur in exceptional cases. On the other hand, pitting is an inherent wear mechanism that, given sufficient time, will eventually develop. Pitting was chosen as the primary wear mechanism of interest for developing this prototype vision-based condition monitoring system due to its inherent nature. For the rest of this paper, only pitting is considered in scope. However, the other wear mechanisms, such as abrasion and adhesion, will be explored in future research.

As of present, the industry has not yet adopted a fully integrated vision system for monitoring gearboxes. However, some research has already been performed on this technology. Xi et al. [18] developed a tunable vision detection platform for online collection of gear pitting images, and developed a Mask R-CNN based measurement methodology for determining the pitting area ratio. They further improved on their research [19] by using an oil baffle-plate to avoid oil splash influencing image quality and investigated using a novel network combining YoloV5 with Deeplabv3+ for real-time detection. Meanwhile, Wang et al. [20] investigated combining a deep convolution generative adversarial network (DCGAN) with a convolution segmentation network (U-net) for measuring the area ratio of gear pitting over several fatigue tests using an automated vision detection device. Miltenovic et al. [21] investigated the possibility of tracking a single, artificially seeded defect in time by using a Deep Faster-CNN model and including a visible number on the side of the gear to determine the specific tooth flank. The mentioned research [18–20] primarily focuses on the detection accuracy of the model and how to improve it further. However, there is still information lacking on the spatiotemporal evolution of gear pitting. Furthermore, all studies [18–21] focused on in-situ visual acquisition and pitting detection of spur gears. Nowadays, helical gears are most commonly used for high-power applications. However, acquiring images on helical gear adds an extra dimension of complexity due to the helical angle. Thus, there is still a need to research automated in-situ image capturing of helical gears, from which relevant metrics can be extracted related to the spatiotemporal evolution of gear pitting.

This paper investigates how gear pitting can be induced and monitored during operation of helical gears, utilising a back-to-back gear tester (section Sect. 2.1). Afterwards, it elaborates on how visual monitoring is achieved by integrating a vision-based condition monitoring system (VCM) (section Sect. 2.2) and the development of an image processing pipeline (section Sect. 2.3) to process the acquired images. This includes grouping the images of gear surfaces by tooth through image correlation, and automated annotation for the extraction of metric on gear pitting. Furthermore, this paper highlights how these metrics can be used to analyse the spatiotemporal evolution of gear pitting for a singular accelerated lifetime test of gear pitting (section Sect. 3). Finally, section Sect. 4 elaborates on the accuracy of the proposed vision system and the practical use of this monitoring approach in the industry.

## 2 Methodology

To generate datasets containing information on the initiation and evolution of pitting on the surface of individual gear teeth, an accelerated lifetime test was performed on an FZG test rig. During this test, automated in-situ image

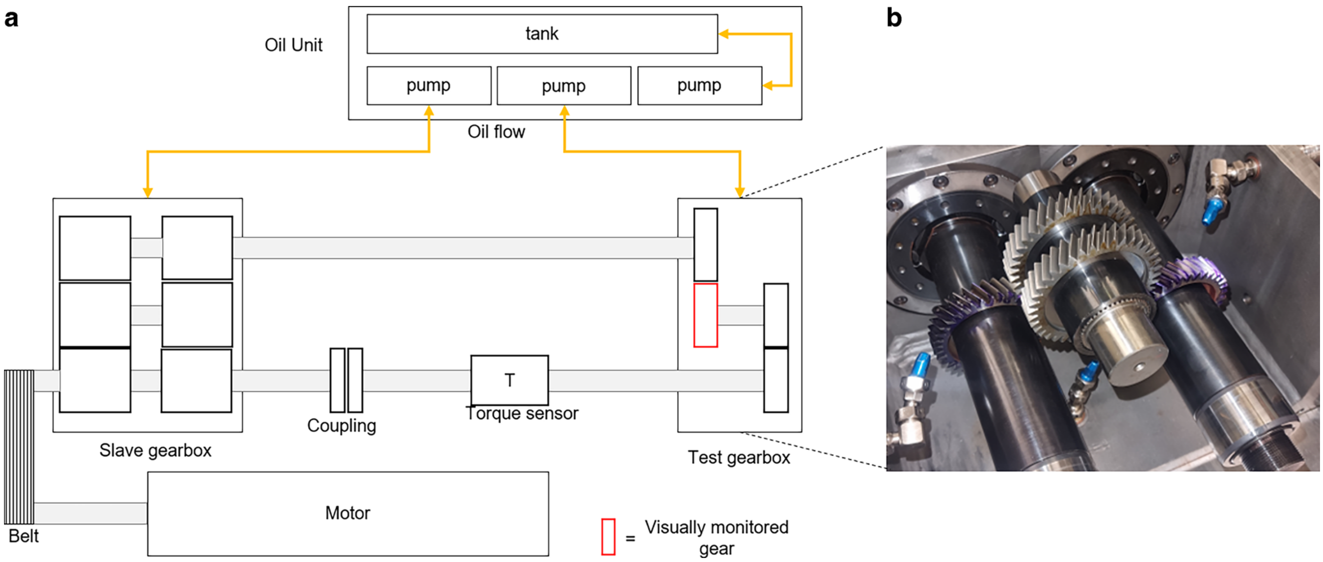



**Fig. 1** FZG multistage gear test rig. In **a** Schematic overview of the main components. In **b** Triaxial design of the test gearbox

acquisition was performed using a novel vision-based condition monitoring system.

## 2.1 Accelerated lifetime testing to generate gear pitting

### 2.1.1 Multistage FZG test rig

A Strama-MPS® multistage FZG (Forschungstelle für Zahnräder und Getriebebau) gear test rig generated the gear wear. Figure 1 shows an overview of the main components of the gear test rig. The test rig is arranged in a back-to-back configuration (Fig. 1a) and uses a triaxial multistage design (Fig. 1b), which enables testing of helical gears. The mechanical load is imposed using a friction disk coupling in which an angular displacement is introduced. The gears operate under spray lubrication with a controlled flow rate.

### 2.1.2 Design of experiments

To design an experiment that facilitates pitting generation and restricts it only to the visually monitored gear (highlighted in Fig. 1a), a gear pair with differing hardness was used. An overview of the gear and pinion specifications can be found in Table 1.

An overview of the test run and lubrication parameters can be found in Table 2. In this table, it can be noted that the torque in the system does not remain constant. In fact, initially applied torque consistently dropped over time

**Table 1** Specifications of the gear and pinion

| Property | Value |
|---|---|
| Gear type | Helical |
| Tangential modulus | 2.1 mm |
| Pressure angle | 20° |
| Helix angle | 23.5° |
| Pinion teeth | 38 |
| Pinion width | 17.10 mm |
| Gear teeth | 41 |
| Gear width | 14.70 mm |
| Tooth height | 6.5 mm |
| Hardness monitored gear | 200 HV |
| Hardness other gears | 550 HV |
| Initial roughness $R_a$ | 0.6 µm |

**Table 2** Pitting accelerated lifetime test specifications

| Property | Value |
|---|---|
| *Test run parameters* | |
| Rotational speed | 2560 rpm |
| Torque | 60–90 Nm |
| Duration | 205 h |
| *Lubrication* | |
| Lubrication system | Recirculating spray lubrication |
| Lubricant | Shell Omala S2 GX 150 |
| Flow rate | 2 L/min |
| Filter size | 10 µm |
| Viscosity @40 °C | 150 $mm^2/s$ |
| Viscosity @100 °C | 14.8 $mm^2/s$ |
| Viscosity index | 97 |
| Density | 892 $kg/m^3$ |
| Temperature | 25–60 °C |

due to fretting occurring between the shaft-gear connection. To avoid excessive decrease of the torque level, an alarm was integrated to the FZG test rig. At 65 Nm, the operator was notified, while at 60 Nm, the test was automatically paused until the load was reapplied. Additionally, the fact that torque is not constant does not pose a problem for this research, as the objective is to explore a prototype VCM and see how it can be used for monitoring and analysing gear pitting. To ensure acceleration of gear pitting, the test run parameters were tuned so that the gears operate under mixed lubrication. This way, the pitting defect is surface initiated due to asperity contact while ensuring sufficient lubrication to minimize other wear mechanisms such as abrasion and keeping the load low enough to prevent adhesive wear. Test parameters were verified by determining the ratio of the composite surface roughness and the film thickness, i.e., $\lambda$-value ($1 < \lambda < 3$ for mixed lubrication), for a range of loads at a fixed rotational speed of 2560 rpm. The $\lambda$-value was estimated based on the theoretical minimal film thickness ($h_{\min}$) through the Moes fit calculation [22]. For a given film thickness $h_{\min}$ and surface roughness of both the gear $R_g$ and pinion $R_p$, the $\lambda$-value can be calculated using Eq. 1. The estimated $\lambda$-value in the function of the applied torque at the start of the test is plotted in Fig. 2. Due to the gear's polishing during the run-in, the gears' composite roughness will be reduced during the first hours of the test, resulting in an increasing $\lambda$ value. Based on previous tests, the roughness $R_a$ of each gear after run-in is estimated to be 0.4 µm, resulting in a $\lambda$ between 1.65–1.80. The total test duration was set to 205 h, resulting in a similar number of rotational cycles (e.g. $28 \cdot 10^6$ cycles) as found in a standardized FZG pitting generation procedure [23].

$$\lambda = \frac{h_{\min}}{\sqrt{R_g^2 + R_p^2}} \tag{1}$$

**Table 3** Specifications of the vision system

| Property | Value |
|---|---|
| Camera | 2 MP IDS 3360CP |
| Lens | RICOH FL-CC3524-5MX |
| Lighting | 7.6 W CCS LDR2 LED ring light |
| Frame rate | 20 fps |
| Image pixel size | 22.4 µm |
| Exposure time | 10 ms |
| Resolution | 2048 × 1088 |

## 2.2 The vision-based condition monitoring system (VCM)

A vision system was integrated into the FZG gearbox to capture the initiation and evolution of pitting on the tooth surface. The specifications of this vision system are shown in Table 3.

A mounting post with 6 degrees of freedom was used to optimally position the camera, as shown in Fig. 3. This mounting post was fixed to the top of the gearbox using a magnet with sufficient pulling force to prevent the camera from moving. Three parameters need to be considered to determine the camera's optimal position. (1) the field of view (FOV), i.e., the plane visible to the camera; (2) the working distance (WD), i.e., the distance at which objects are in focus, which depends on the focal length of the lens and magnification; (3) the depth of field (DOF), i.e., the

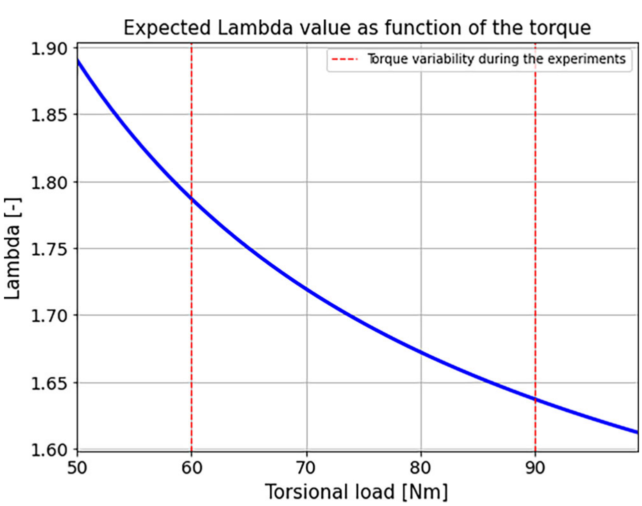



**Fig. 2** Calculated $\lambda$ value as function of the torque

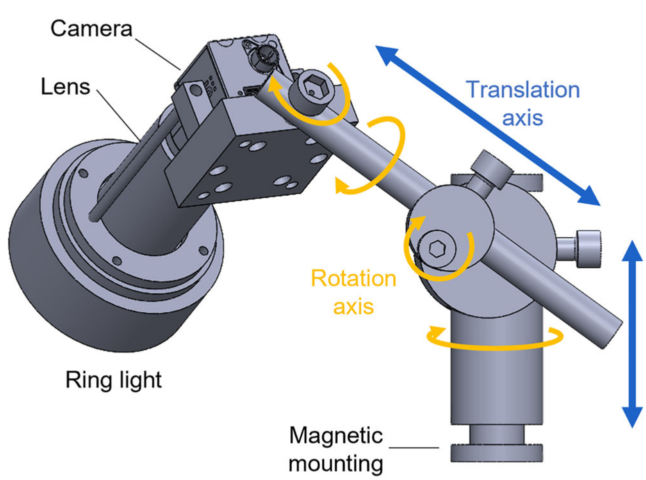



**Fig. 3** Overview of degrees of freedom of the magnetic mounting post

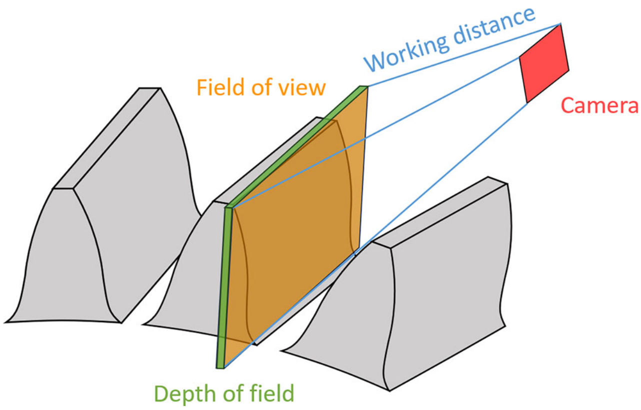



**Fig. 4** The field of view, working distance and depth of view of a vision system

range at which objects appear sharp, starting from the WD. These parameters have been visualised in Fig. 4. The camera needs to be positioned to balance these three parameters.

The lens utilized in this research has an adjustable WD, between 100 mm and "infinity", and an adjustable aperture to increase/reduce the depth of field. In practice, the camera was positioned according to the following procedure:

1. Position the camera perpendicularly to the tooth flanks using the magnetic mounting post.
2. Set the depth of field (DOF) as high as possible by adjusting the lens aperture. The available light and set exposure time (10 ms) limit the maximum DOF, as closing the aperture ring increases the DOF but the system requires more light to compensate the exposure.
3. Reduce the distance between the camera and the tooth flank whilst ensuring the whole tooth flank stays within the DOF.
4. Adjust the working distance of the lens for optimal focus.

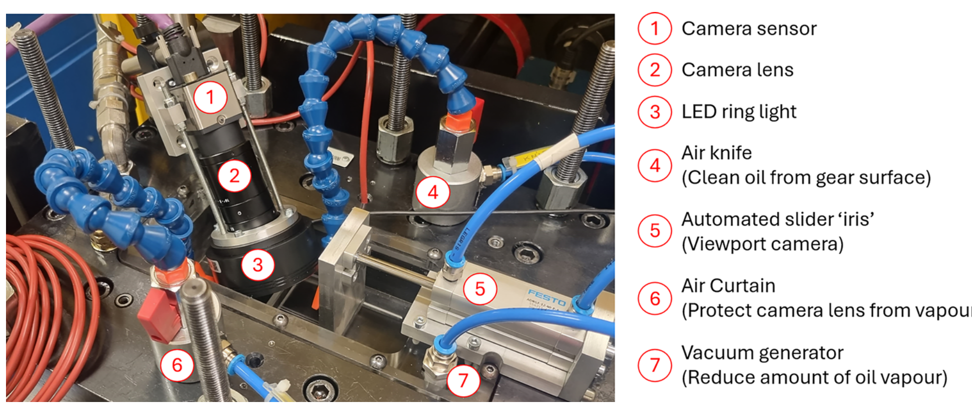
This procedure was performed once at the start of the test, after which the camera did not need to be repositioned.

To ensure the quality of the images, an automated oil splash mitigation strategy was conceptualised, developed and tested to circumvent the influence of the lubricant on image quality. The vision system with the auxiliary components used for this oil splash mitigation strategy is highlighted in Fig. 5.

The following steps are associated with the oil splash mitigation strategy:

1. During image acquisition, the oil flow towards the test gears is interrupted, and the speed of the test rig is reduced to 1 rpm. This is done to minimise the effect of the rotating gears whipping the oil. Additionally, this slowdown is necessary to bring the rotational speed below the set acquisition speed of the camera (20 fps), which is lower than the maximum possible frame rate of the camera (153 fps) because of the selected exposure time (10 ms), as higher exposure times prevents acquiring sharp images of the gear flank.

**Fig. 5** Overview of the VCM-system's components



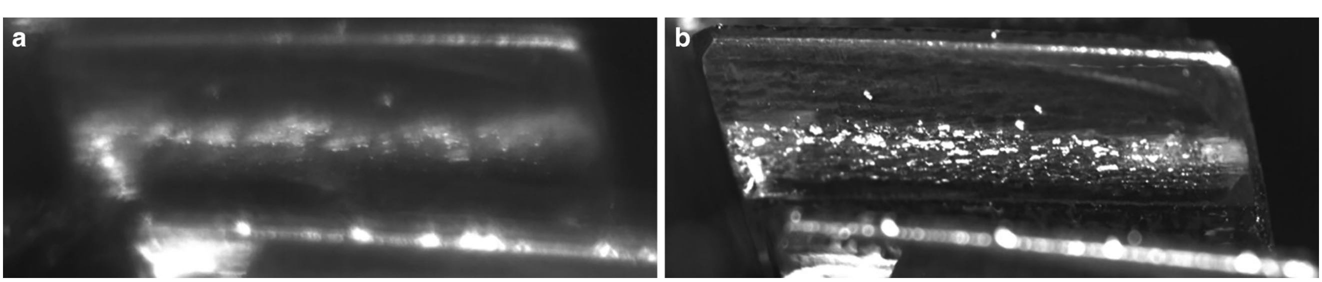




**Fig. 6** Difference between contaminated (**a**) and clean (**b**) lens during image acquisition

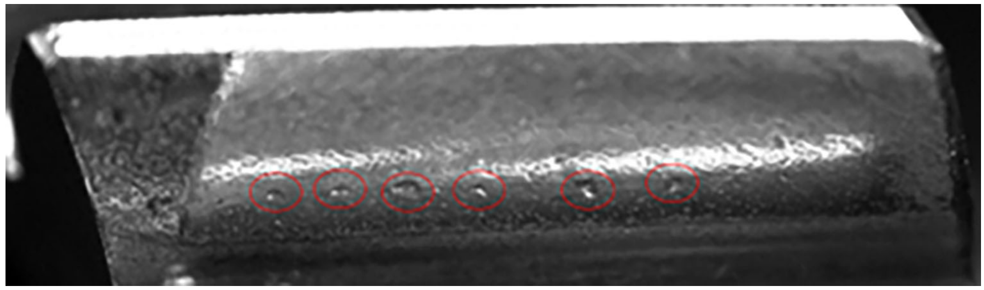


**Fig. 7** Fake defects due to oil residue left on the tooth

**Table 4** Properties of the image dataset

| Property | Value |
|---|---|
| Colour mode | Gray-scale |
| Captures frames | 5248 (total)—128 (per tooth) |
| Cropped size | 896 × 320 |

**Table 5** Percentual cost of hardware components for the VCM

| Hardware | Percentual cost (in %) |
|---|---|
| Camera | 30 |
| Lens | 10 |
| Lighting | 25 |
| Connection cables | 5 |
| Auxiliary system | 30 |

2. A vacuum pump (item 7 in Fig. 5) extracts as much hot oil vapours as possible from the gearbox. These vapours could otherwise contaminate the lens during the next step of opening the slider window. The effect of the lens contamination can be seen in Fig. 6.
3. An automated pneumatic slider window (item 5 in Fig. 5), which protects the camera during nominal operation, is opened while acquiring the images. This slider window functions as an iris for the camera. If this were not done, an oil build-up would be left on the plexiglass window, which can remain there for hours to days, obstructing clear images.
4. A pressured air nozzle (item 6 in Fig. 5) is activated to create an air curtain between the camera lens and the opening in the plexiglass. This air curtain forms a second barrier to protect the lens from getting contaminated due to remaining oil vapour escaping from the gearbox.
5. A pressured air nozzle (item 4 in Fig. 5) is aimed at the gear surface, functioning as an air knife to momentarily clean the gear surface before the vision system captures it. Oil residue that is left on the surface can have a similar appearance to surface pitting. By utilising this air knife, these "fake defects" are avoided. Oil droplets resembling surface pitting are illustrated in Fig. 7.

Once the images have been acquired, which takes 1 min, the oil splash mitigation strategy is undone so the machine can return to its nominal operation. A set of images was acquired every 30 min, and the entire process was automated through Python. The specifications of the image dataset generated during the 205 h accelerated lifetime test can be found in Table 4.

The total hardware cost of the entire visual monitoring solution was a few thousand euros (percentual costs shown in Table 5). However, two important aspects related to the total cost should be noted. Firstly, the integration is a complex and time-consuming task, meaning development costs will outweigh hardware costs. Secondly, if there is a need for a high frame-rate camera (e.g. 2000 fps), the total hardware costs are estimated to be 5–10× larger, depending on specific needs.

### 2.3 Image processing pipeline

An additional workflow is required to extract quantitative results from the in-situ images. This workflow will be referred to as the image processing pipeline, and each step will be further explained in this section. Figure 8 shows an overview of the image processing pipeline. An overview of the workflow is found below.

1. Capturing an image burst for a full gear rotation.
2. Filtering the image burst down to 1 image per tooth.
3. Cropping the image so only a single tooth remains.
4. Group the images by individual tooth.
5. Iterative annotation process to generate the training dataset for the deep learning model.
6. Annotation of the entire dataset by trained deep learning network.

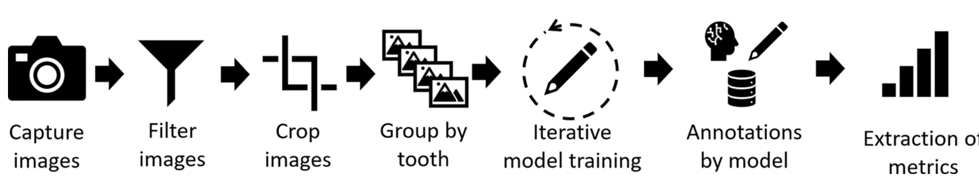



**Fig. 8** Flow diagram of the image processing pipeline

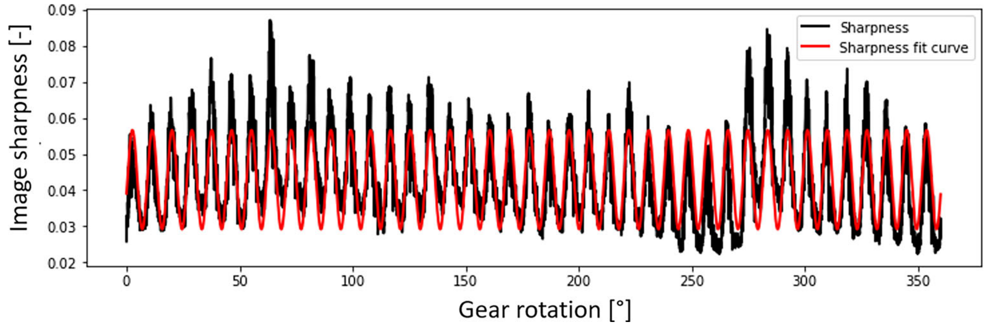



**Fig. 9** Sharpness of image burst (top) and sine fitting (below), in function of gear rotation [°]

7. Extraction of the quantitative metrics from the annotated images and conversion to metric units through calculating the pixel size.

During the acquisition of the images, a large burst of images is taken for one full gear rotation. This is done since determining the perfect position of the gear when the defects are in focus is not feasible when utilising this test rig, as it is incapable of persistently monitoring the position of the visually monitored gear. As a solution, a burst of an entire rotation is captured, after which this burst is filtered so that only a single image remains for every tooth. This filtering is done by calculating a sharpness score [24] for each frame in the image burst. Due to the cyclic nature of the sharpness, as the gears move in and out of focus with a constant speed, a sine function can be fitted over the sharpness burst, as shown in Fig. 9. This sine fitting is then used to extract an image for every tooth. This process is done automated and performed immediately after capturing the images to avoid saving a large amount of unnecessary data, as this process reduces the 1350 frames captured during the burst to a mere 41 frames when filtered, matching with the number of tooth of the visually monitored gear.

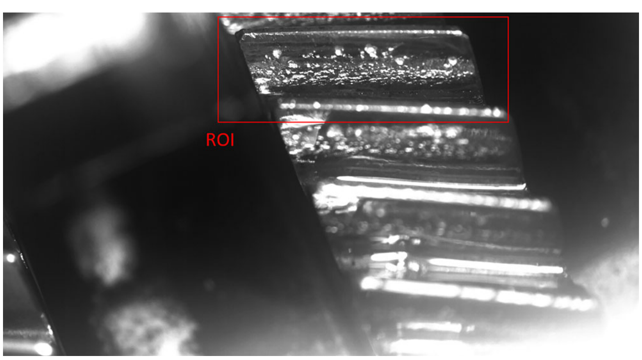



**Fig. 10** Uncropped image captured during the accelerated lifetime test, region of interest (ROI) is indicated

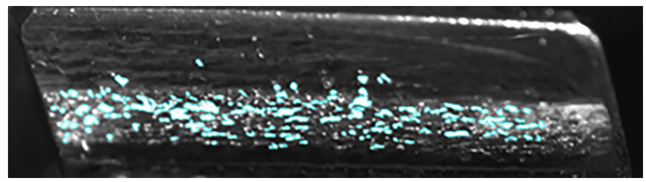


**Fig. 12** Annotation of pitting observed on a gear tooth

Once all the images have been filtered, they are cropped so that only the tooth in focus remains, removing irrelevant information and reducing the data size. An uncropped image is shown in Fig. 10.

After filtering, all images are sorted by tooth using an image correlation algorithm. Each tooth surface contains multiple unique visual identifiers (e.g., machining marks and pitting defects) behaving as a visual fingerprint as shown in Fig. 11. This visual fingerprint allows assigning each image to a specific tooth number (from 1 to 41). Additionally, considering that teeth always appear in the same order, a good correlation of a certain image to its previous (already assigned to a certain tooth) is enough to assign tooth numbers to all images on the same burst. Additional matched images further confirm the correlation. The effectiveness of this algorithm was verified by manual inspection.

Afterwards, all images need to be annotated to extract metrics from the images, as shown in Fig. 12. These annotations can then be processed to extract metrics. However, the annotation process is slow and labour-intensive, sometimes taking up to 1 h to annotate a single image. Since this dataset contains over 5000 images, annotating it would take over one full year.

Deep learning can be leveraged to speed up the annotation process. A convolution neural network is used for

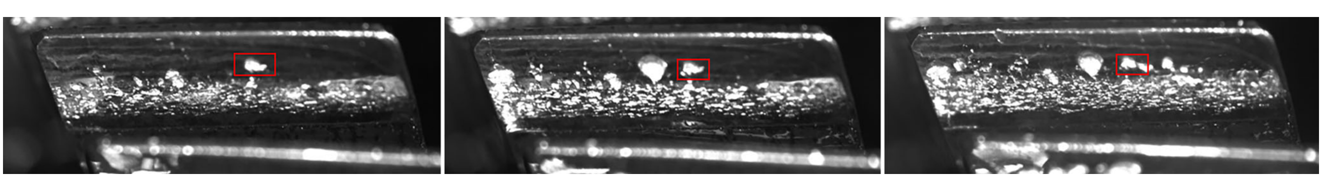


**Fig. 11** Example of a visual identifier on a gear tooth

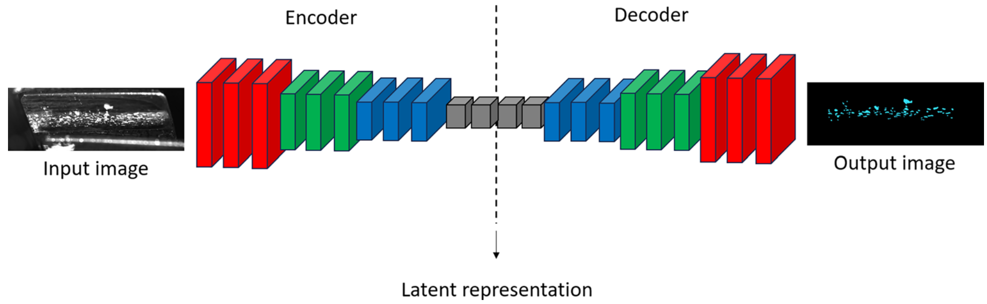



**Fig. 13** Working principle of the encoder-decoder deep learning network

this automatic annotation of gear teeth, which utilises the encoder-decoder principle illustrated in Fig. 13. This network takes an image and encodes the information until it's left with the latent representation. This latent representation is a condensed vector that only contains essential information. Afterwards, the latent representation is decoded into an image containing only the vital information, the pitting quantity. The network used in this paper was the Feature Pyramid Network architecture [25], using the pre-trained weights for the imagenet dataset [26]. The quantity of necessary training data can be significantly reduced by starting from a pre-trained network that already contains many notions on the underlying features of images [27]. This is because images contain many standard features, such as basic geometrical shapes (e.g., squares, circles, and triangles).

However, a small subset of the entire dataset must be annotated to train this deep-learning model. An iterative training strategy has been devised to reduce annotation effort, as shown in Fig. 14.. A first set of key frames for annotation are selected through an active learning loop, which determines the images that will provide the most information to the deep learning model [28]. Next, people trained to recognize pitting annotate these key frames using the open-source CVAT tool [29]. All defects with a minimum size of 10 pixels in horizontal or vertical directions are annotated in these key frames. This 10-pixel limit is set, as annotating anything below this threshold becomes too tricky as the pit boundaries are unclear due to the low resolution. Once all key frames are annotated, the model is trained for the first time. Using the trained deep learning model, the entire dataset is annotated. To determine whether the annotations are sufficiently accurate, metrics are extracted from the annotations by the deep learning model and the curves provided by these metrics are analysed. As the evolution of these metrics is expected to be a smooth increase over time, this can form the first basis for the model's accuracy. Additionally, the annotations can be mapped on the correlating image for a more in-depth accuracy analysis. If the model's accuracy is insufficient, a second set of key frames is determined using active learning. Furthermore, the annotations by the model can be uploaded onto the CVAT tool, which the expert can correct. The annotation time per image can be further reduced by correcting the model predictions (instead of annotation from scratch). This iterative loop is performed until the model reaches the desired accuracy or the end-users' satisfaction.

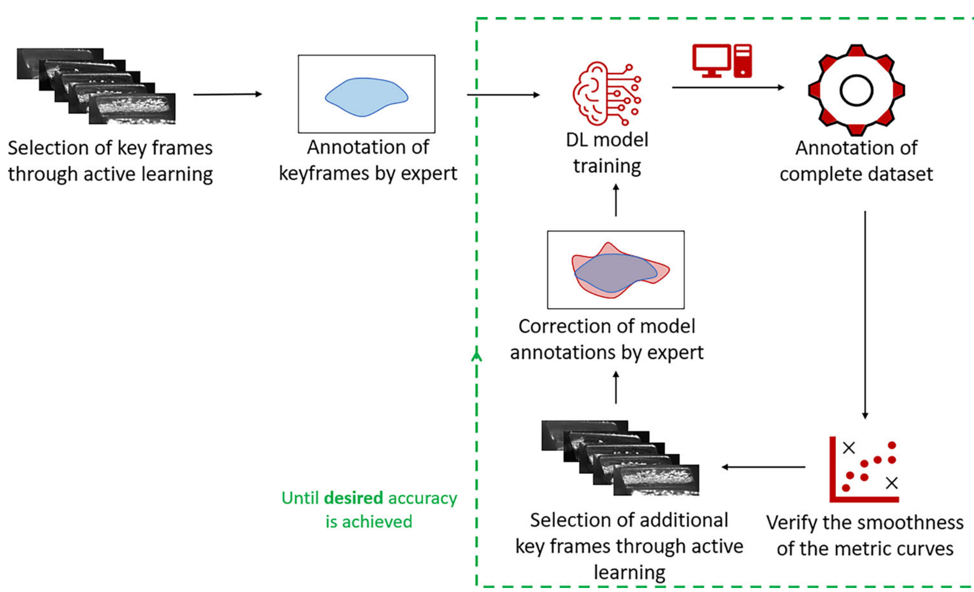



**Fig. 14** Iterative annotation strategy for training the deep learning model

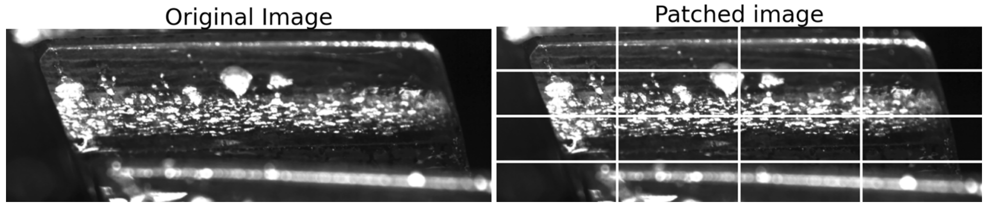



**Fig. 15** Example of a patched image for upsampling the training dataset

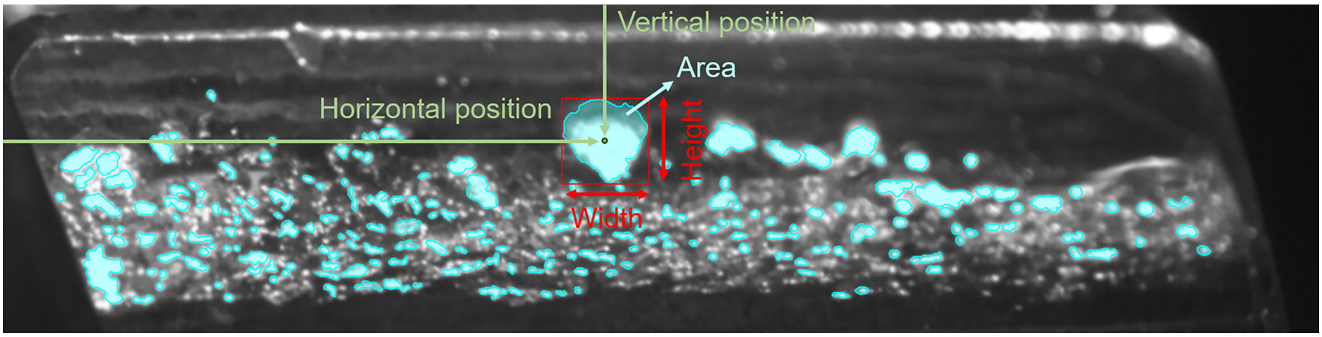



**Fig. 16** Relevant metrics that can be extracted from the annotations

For this dataset, 14 images (0.3% of the full dataset) were annotated to train the model. Before training the model, upsampling of the training dataset is performed by patchifying the training images. This approach increases the training pool by approximately a factor of ten. An example of a patched image is shown in Fig. 15.

Once the dataset has been fully annotated, the individual pits can be processed to extract relevant metrics, as shown in Fig. 16. Metrics that can be considered relevant are:

- Area of the pit(s)
- Amount of pits
- Height of the pit(s)
- Width of the pit(s)
- Horizontal position of the pit(s)
- Vertical position of the pit(s)

Furthermore, statistical indicators, such as the mean, maximum, and minimum, can be calculated based on the annotations.

Finally, to convert these metrics into metric units, the pixel size is required. The pixel size depends on the camera, lens, gear geometry and the position of the vision system. As the width of a tooth is known in both millimetres and pixels, the pixel size $\Delta x_p$ can be calculated using Eq. 2.

$$\Delta x_p = \frac{w_{f,\mathrm{px}}}{w_{f,\mathrm{mm}}} = \frac{713\,\mathrm{px}}{16.0\,\mathrm{mm}} = 22.4\,\frac{\mu\mathrm{m}}{\mathrm{mm}} \tag{2}$$

The face width $w_f$ in pixels can be measured directly from the images and is equal to 713 px. As the camera is perpendicular to the tooth face, The face width $w_f$ in mm will differ from the gear width $w_g$ due to the helix angle $\beta$. This face width $w_f$ can be calculated using the geometry in Table 2 and Eq. 3, resulting in a pixel size of 22.4 $\frac{\mu\mathrm{m}}{mm}$.

$$w_f = \frac{w_g}{\cos\beta} = \frac{14.7\,\mathrm{mm}}{\cos 30°} = 16.0\,\mathrm{mm} \tag{3}$$

## 3 Results

This section analyses the spatiotemporal evolution of a single accelerated lifetime test for gear pitting based on the metrics extracted after applying the image processing pipeline described in Section 2.3.

### 3.1 Area of the pits

When considering area, the absolute area $A_{\mathrm{abs}}$ in mm$^2$, and the relative area $A_{\mathrm{rel}}$ in % can be considered, where the relative area is calculated using Eq. 4, using the face width $w_f$ and tooth height $h_t$. Further exploration of the evolution of the total pitted area will be based on the relative pitting area. This evolution for different gear teeth is shown in Fig. 17. In this figure, four different zones are indicated by the red dashed lines, each representing a different stage in the health of the gear.

$$A_{\mathrm{rel}} = \frac{A_{\mathrm{abs}}}{w_f h_t} \tag{4}$$

Before rapidly transitioning between zones, a temporary state of equilibrium is reached, and a new underlying mechanism causes damage to further progress. Images correlated to each zone are shown in Fig. 18.

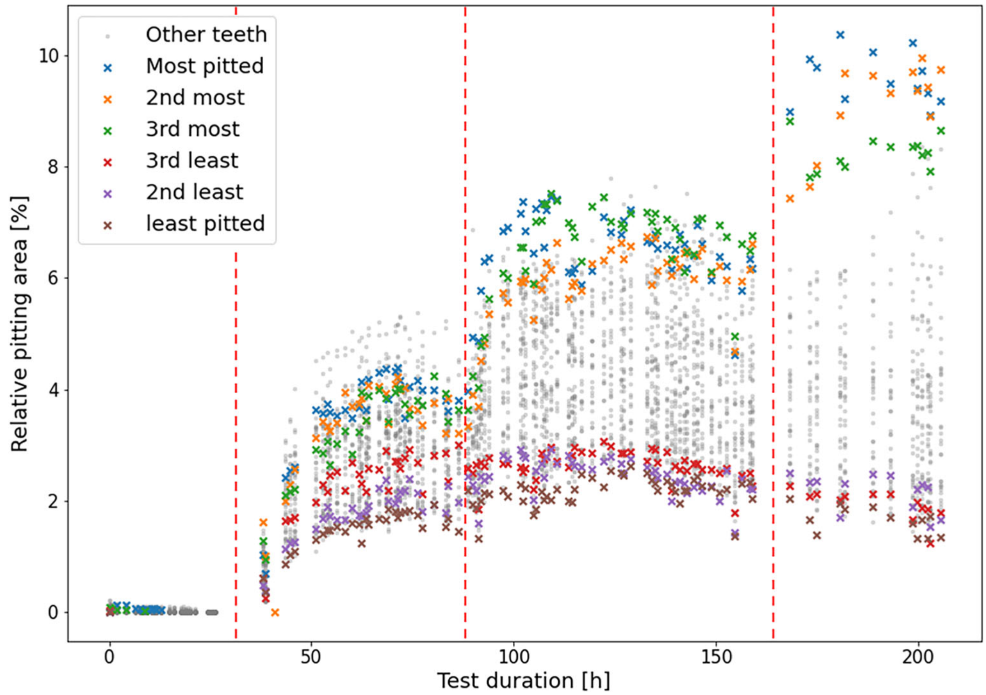



**Fig. 17** Evolution of the relative pitted area during the accelerated lifetime test (total flank area: $104\,mm^2$)

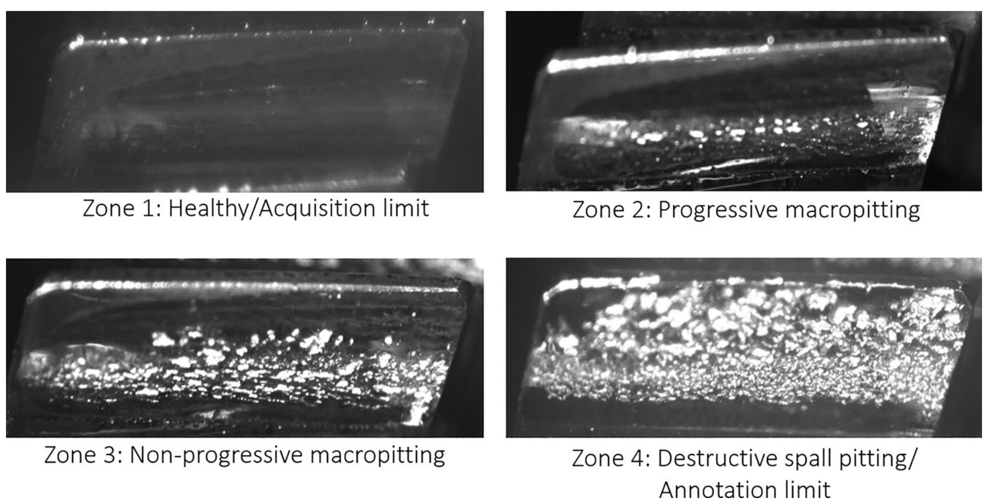



**Fig. 18** Images correlated to the four different zones

- **Zone 1—Healthy/Acquisition limit:**
  Here, the gear is still healthy, or the damage is below the minimal set defect size for annotation. The minimum set size for annotation was 10 pixels (230 µm). This first zone was found for the first 40 h of the test. However, some noise can be detected in this zone, resulting from wrong annotations by the deep learning model, which will be further explored in section Sect. 4.1. Therefore, further results will be explored from this point forward to avoid the influence of this noise.
- **Zone 2—Progressive macropitting:**
  At this stage, pits gradually appear along a consistent horizontal range on the gear surface, further described in section Sect. 3.3. These pits correspond to progressive macropits as described by the ISO 10825:2021 standard [13] and result from the meshing of both gears together.
- **Zone 3—Non-progressive macropitting:**
  In this zone, macropits appear that differ from the previous progressive macropitting as they have a large area, are more circular, appear more sudden, and are found higher on the tooth flank. These pits matches the description of non-progressive macropits in ISO 10825:2021 standard [13]. They result from wear debris getting caught between the meshing gears, creating indentations that cause local high-stress concentrations during meshing.

- **Zone 4—Destructive spall pitting/Annotation limit:** At this point, the progressive and non-progressive macro-pits start to coalesce, creating much larger pits with an irregular shape over the entire gear surface. These pits correspond to spall pitting as described by the ISO 10825:2021 standard [13]. The additional wear debris resulting from this spall pitting causes some teeth to show extreme pitting (e.g. the blue, orange and green trends), which the deep-learning model severely underestimates. Both these effects result in the introduction of an annotation limit, as it becomes increasingly tricky to recognize clear borders of the pits. Both trained annotators and the deep learning model experience this limit.

Additionally, the mean value of the pitted area was investigated. However, it was found that it could not be correlated to the damage progression due to gear pitting, as it showed no trending in zones 3 and 4, unlike the graphs for the individual teeth. This is due to the mean value only giving a global view on the gear health, while from a monitoring perspective it's commonly the most pitted teeth that supply the most information on gear health and will determine end-of-life. The inability to correlate mean pitted area with damage progression, specifically in zones 3 and 4, further illustrates the need for a methodology that allows for analysis of the individual teeth. Because of the reason mentioned above, the mean value was not included in Fig. 17.

Although the pitting area is physically expected only to increase in size, observe the fluctuations in the pitted area curve for the different teeth in Fig. 17. Furthermore, in zone 4, some curves even suggest an overall reduction in area. Although it is possible there are minor reductions in the pitted area due to plastic deformation at the edges of the pits or by wear of the gear profile. The model's inaccuracy, especially in zone 4, was considered to be the main driving factor to explain these fluctuations and the downward trend. The accuracy of the model was found to be influenced by three aspects. An in-depth discussion on the accuracy of the model will occur in section Sect. 4.

Similar trends to the area evolution can be observed in the following metrics: amount of pits, total pitted height and total pitted width, as illustrated in Fig. 19. Section 3.2 will investigate the correlation between the area and the height and width.

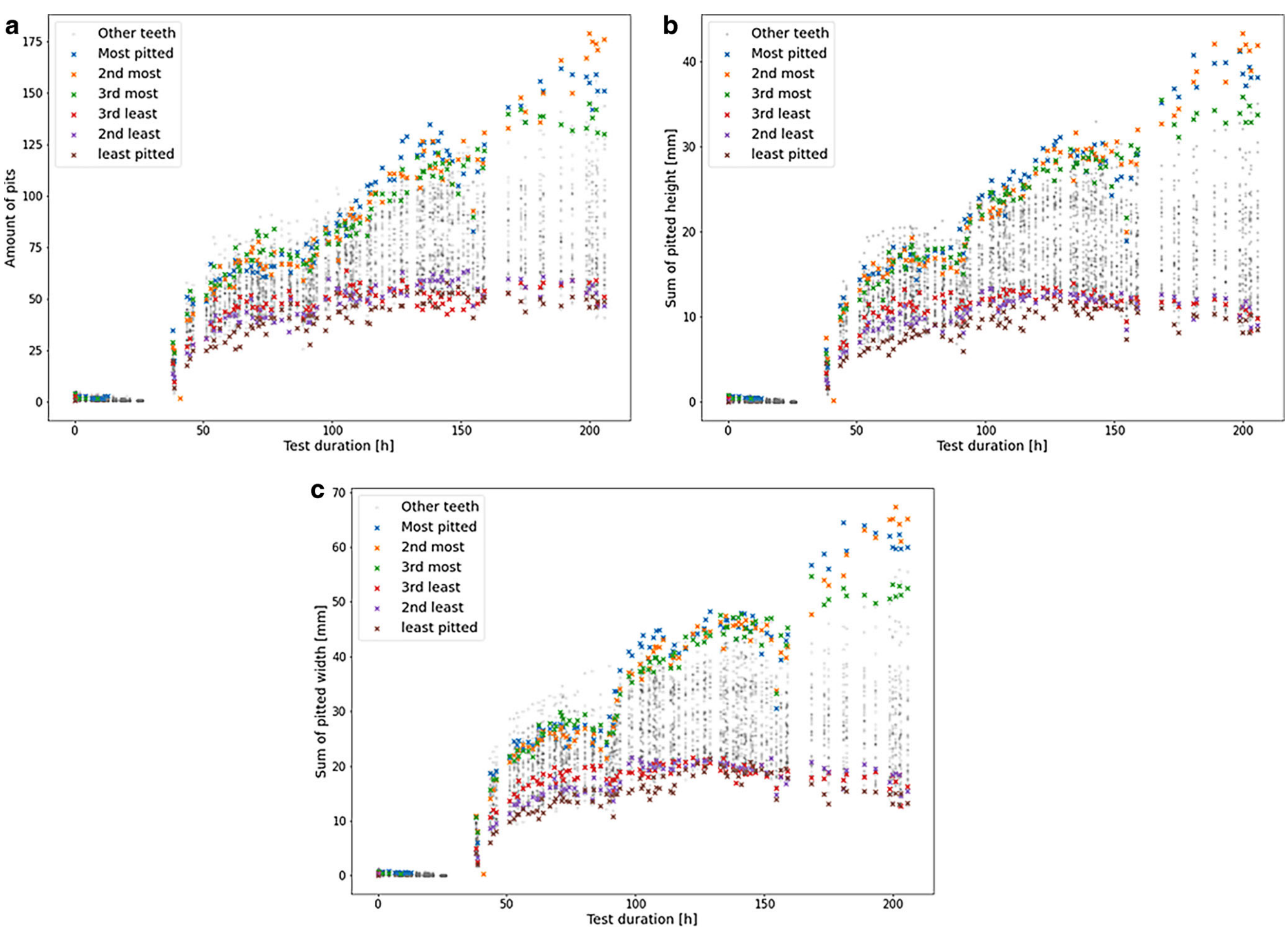



**Fig. 19** Evolution of the number of pits (**a**), summed pit height (**b**) and summed pit width (**c**) during the accelerated lifetime test

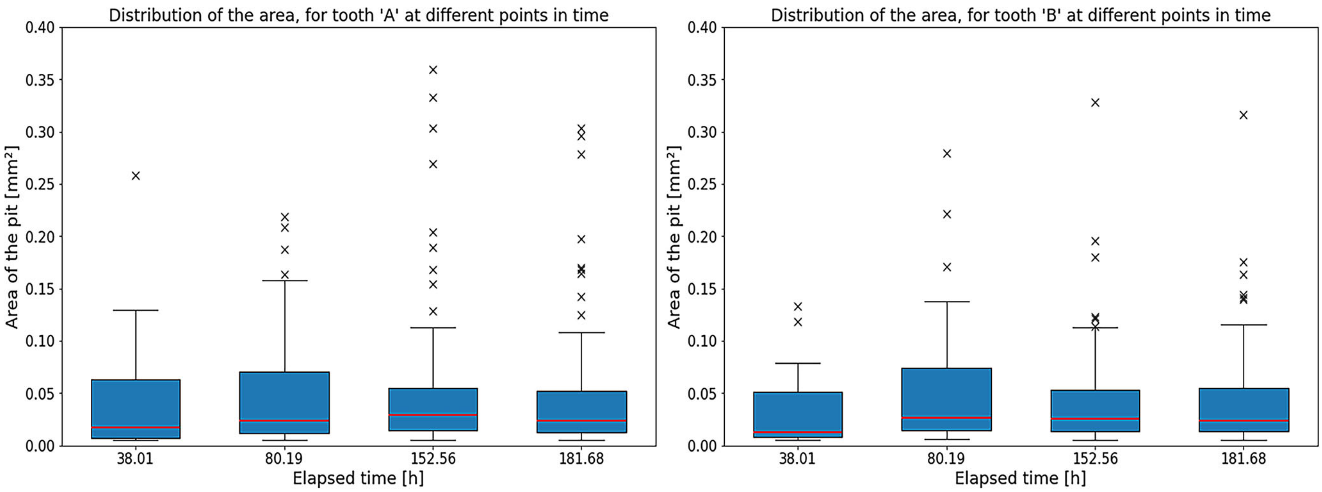



**Fig. 20** Box plot showing the distribution of the area of the pits at four different points in time, for two different teeth

Furthermore, the distribution of the pit areas can be investigated when observing a singular tooth. In Fig. 20, box plots indicate the area's distribution at four different points in time for two teeth. For visualisation purposes, the maximum area is limited to 0.4 mm², which encompasses 93% of all pits found at the end of the test. Two observations can be made. Firstly, the distribution over time remains primarily unaffected, and mainly more large pits appear (x-marks in Fig. 20). Secondly, the distribution of the pitted area is similar between different teeth, where, in general, the most significant difference lies in the large pits.

## 3.2 Correlation between the area and the height and width of the pits

This correlation between the total area, height and width of the pits can be further explored by analysing the ratio $r$ of the height and weight for pits as described by Eq. 5. This ratio was calculated for each pit, and afterwards, the average height-to-width ratio was calculated for each tooth, resulting in Fig. 21. At the beginning of the test, there appears to be a larger spread of this ratio caused by the strong influence of outliers due to the small number of pits. However, as the test progresses, this ratio becomes more constant over all teeth. Additionally, it can be noted that on average, this

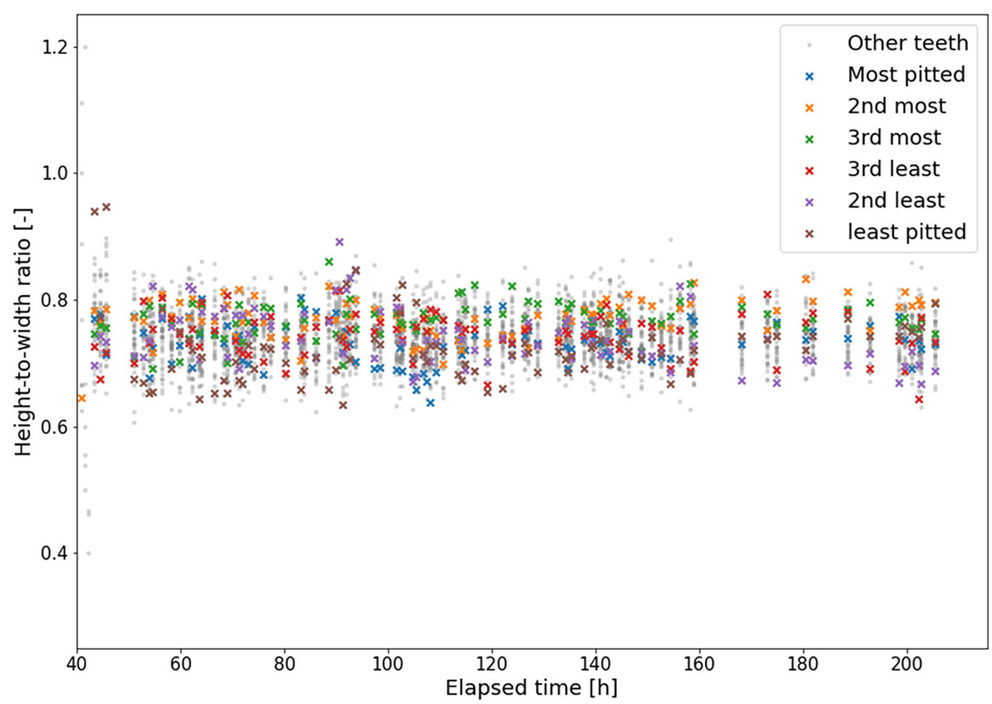



**Fig. 21** Evolution of the average height-to-width ratio for each tooth during the accelerated lifetime test

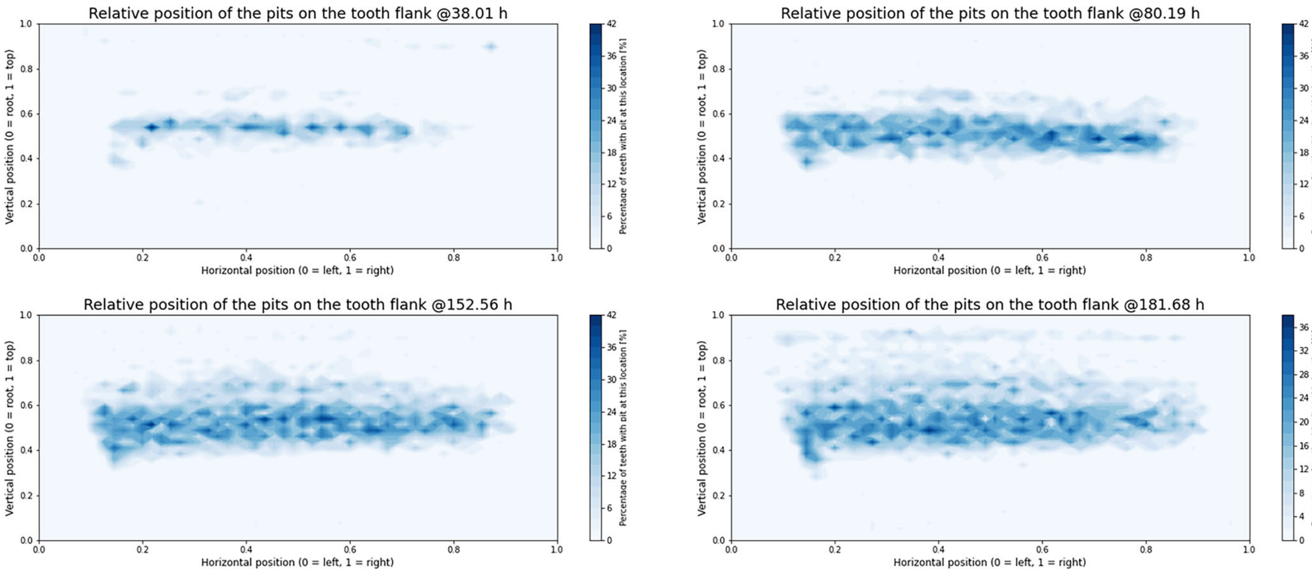



**Fig. 22** Heat map of the likelihood of the locations at which pits were found over all the gear teeth at different points in time

ratio is equal to 77%%, meaning the width of the pits $w_{\mathrm{pit}}$ is about 1.3 larger than their height $h_{\mathrm{pit}}$. Previous research from the same research group [30] corroborates this finding, as pitting was found to develop perpendicular to the rolling-sliding motion direction. Likewise, investigations in literature [31, 32] about micropitting found similar results, as they mentioned that micropitting develops further in the gear axial direction than in the radial direction.

$$r_{\mathrm{pit}} = \frac{h_{\mathrm{pit}}}{w_{\mathrm{pit}}} \tag{5}$$

### 3.3 Position of the pits on the gear flank

The final metric on gear pitting that will be explored in this paper is the position of the centre of the pits, precisely their combined horizontal and vertical position on the tooth surface. To visualise this, heat maps of the location of the pits over all teeth at different points in time are shown in Fig. 22. At the beginning of the test (38.01 h), most pits appear around the same vertical range (0.4–0.6), and develop higher on the tooth flank due to wear debris as the test progresses.

## 4 Discussion

Firstly, the accuracy of the vision solution will be evaluated. Secondly, the benefit of tracking the evolution of individual teeth will be discussed. Finally, the usability of such a system in the industry will be discussed, followed by an elaboration of the use of a vision solution in relation to the monitoring techniques currently used.

### 4.1 Evaluation of the accuracy of the vision solution

The accuracy of the vision solution can be validated by comparing the annotations generated by the deep learning model with the manual annotations performed by trained annotators. These manual annotations can be assumed as baseline due to the following factors:

- In practice, manual investigation (often only qualitative) of the gear surface is conducted through vision inspection, for example, using a borescope.
- The image quality is assumed to be sufficiently high to represent the gear surface's condition accurately.
- To minimize errors during annotation, frames of the same gear flank captured at different timestamps were analysed.
- While some degree of error in manual annotations is inevitable due to the subjectivity of the annotator, this error is minimized by the annotation guidelines discussed in Sect. 2.3.

Initially, the accuracy of the model was evaluated qualitatively, and the following three factors were found to be the leading factors influencing the model's accuracy:

1. Usage of a limited dataset for training the model. The largest possible training dataset (i.e. the full dataset) would be required to reduce this effect as much as possible, eliminating the need for the deep learning model.
2. Bias in the annotation provided by the trained evaluators, this phenomenon is also known as the inter-annotator agreement and is shown in Fig. 23. This figure results from another dataset in which three tribological experts annotated pitting defects. Since then, annotator guide-

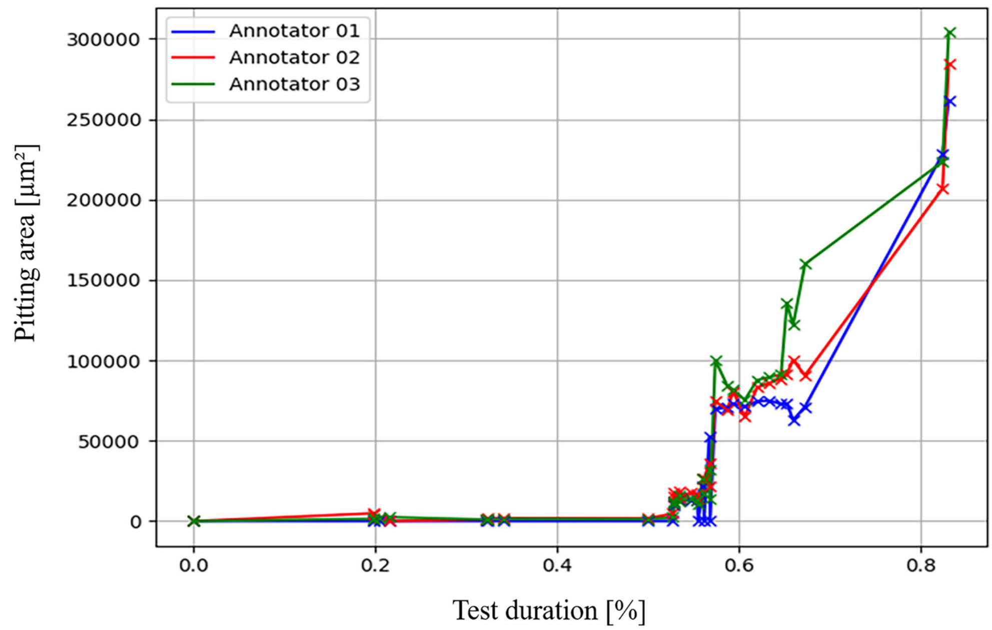



**Fig. 23** Bias in the annotations provided by three different annotators

lines have been devised to reduce this annotation bias. The following guidelines were followed during the annotation of the training dataset: (1) Only defects with a minimal size of 10 pixels are annotated; (2) When two pits are coalescing, the outer boundary will be annotated; (3) Frames adjacent in time must be analysed to avoid wrong annotations, such as caused by lighting effects.

3. Temporary effects on the surface cause incorrect annotations by the model (e.g., reflective patterns due to oil residue). An example of this is shown in Fig. 24; here, two images of consecutive acquisitions are shown where such a temporary effect has been highlighted.

Furthermore, as mentioned in section Sect. 3.1, some noise is detectable during the first 40 h in Fig. 17. Wrong annotations by the deep-learning model cause this noise. As the model is trained to “detect” damage, it struggles when an image is presented without pitting. In this instance, it then tried to annotate a reflective pattern on the top right corner of the gear, which is not actual damage. An example of this is shown in Fig. 25. As the test progressed and more pits appeared, the model was less prone to annotating this

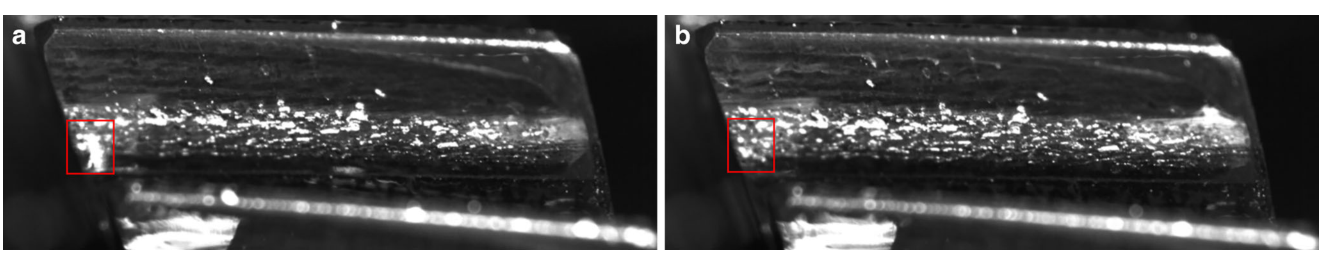



**Fig. 24** An example of a temporary effect, lighting in this case, found in image (**a**) but not visible anymore in image (**b**), which was captured later

**Fig. 25** Wrong annotation of a reflective pattern by the deep-learning model

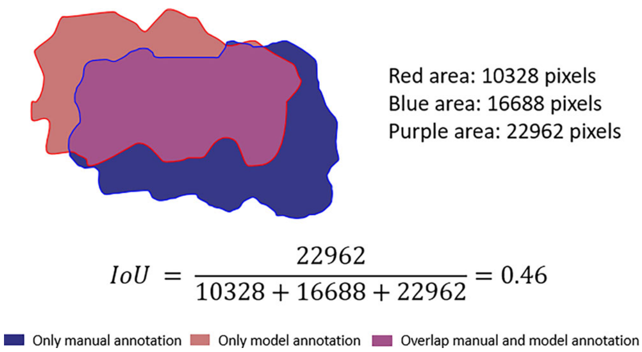



**Fig. 26** Example of Intersection of Union (IoU) calculation for a single pit

reflective pattern at the top of the gear. Due to the small area and small number of wrong annotations, the influence on the area evolution is minor. However, their influence in the early stage of the test is much higher for the other metrics, such as the position. Due to this, some metrics will only be analysed from the 40-hour mark. In the future, this noise can be reduced by:

1. Only providing images to the model that contain damage, either through first manual sorting of the dataset or by using a classification model first to determine whether pitting is present.
2. Detection of the tooth flank and only annotating inside this region. This way, the reflective pattern found at the top of the gear tooth would not be included in the area that requires annotation.

Hereafter, the model's accuracy was evaluated quantitatively through the Intersection over Union (IoU). A calculation example is given in Fig. 26 for a singular pit. An IoU range of 0.50–0.60 for the full training dataset was observed, with a median value of 0.56. Figure 27 illustrates the comparison between manual annotations (blue) and those generated by the deep learning model (red). The model struggles with accurately identifying pit boundaries, leading to underestimating individual pit sizes. Additionally, the model sometimes fails to annotate the smallest pits. Furthermore, the deep learning model occasionally annotates temporary effects (as discussed in the qualitative analysis of the model's accuracy), which are identified and consequently not annotated by the trained annotators, further decreasing the IoU. The values found for the IoU, combined with the results discussed in the quantitative analysis above, indicate that the exact values produced by the model cannot be blindly trusted. However, the general trends in the evolution of the annotations are sufficiently reliable.

## 4.2 The use of vision monitoring in practice

When developing the proposed visual monitoring solution, significant steps have been taken to introduce this novel technique to the industry. However, at this time, this technique is primarily intended for larger gearboxes in remote and/or critical applications. This limitation is due to two factors. The first limiting factor is related to the cost of this monitoring solution (see section Sect. 2.2). The second factor is the requirement for the monitored gearbox to be sufficiently large to accommodate the integration of the vision system. At this time, integrating a vision system into gearboxes would require case-specific solutions. Nevertheless, it was demonstrated in this study the feasibility of integrating such vision system in a test gearbox that replicates the functionality of gearboxes applied elsewhere. For instance, an inspection port could be modified to house a vision system. Additionally, the vision pipeline (Sect. 2.3) could be adapted and used with alternative vision systems. One significant technological challenge remains: preventing contamination of the camera lens by hot oil vapours escaping from the gearbox. Ongoing research is focused on addressing this issue to improve the reliability of the vision system. Aside from this, future research should focus on integrating this vision system in a planetary gearbox, as these are commonly used in the wind industry. This tran-

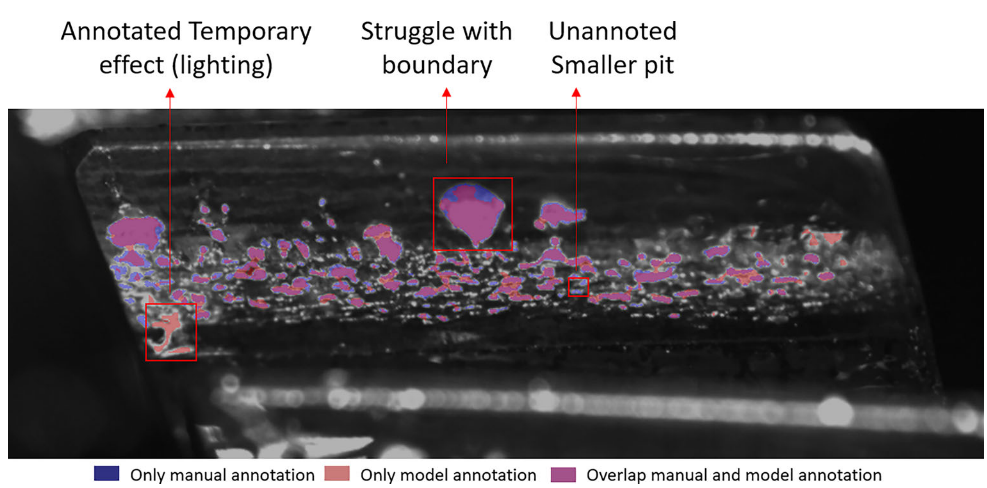



**Fig. 27** Comparison of the manual annotations (red) and those performed by the deep learning model (blue)

sition will bring forward further technical challenges due to the spatial constraints of a planetary gearbox. However, the authors still consider integrating a VCM in a planetary gearbox feasible, as this monitoring system would focus on the output gear of the planetary stage, which contains stationary bearings. The output stage is chosen because pitting is a fatigue phenomenon, so the increased amount of loading cycles will accelerate pitting wear.

A visual monitoring approach has advantages and disadvantages compared to traditional condition monitoring techniques for gearboxes, such as vibrational analysis. The disadvantages of this approach include its higher cost compared to existing alternatives and the need for a case-by-case integration solution. However, given that this technique is designed for larger transmissions in remote or critical applications, the increased cost and integration challenges can be justified. Another limitation is that monitoring every gear within the gearbox may not be feasible due to dimensional constraints. In such cases, the vision solution can focus on monitoring the most critical gear inside the gearbox. On the other hand, the vision system offers several advantages. It can detect damage earlier than alternative techniques, directly confirm the presence of damage on the gear surface, and provide both quantitative and qualitative information about the damage. Furthermore, the ability of a VCM to differentiate between the different gear teeth further improves its monitoring capability over traditional techniques. While traditional techniques commonly only indicate the overall health of the entire gear, the proposed solution provides the degradation curve per tooth. This allows various industries to utilize their own thresholds for when gear is considered end of life due to pitting based on the indicators extracted, as what is considered end-of-life is highly application dependent.

## 5 Conclusion

This paper investigated the use of a vision-based condition monitoring (VCM) system for the in-situ monitoring of gears. A VCM was successfully integrated into the back-to-back gear tester to automatically capture images of the gear flanks during an accelerated lifetime test for gear pitting. An oil splash mitigation strategy was developed and implemented to avoid oil splash hindering the quality of the pictures. Afterwards, this paper elaborated on the developed image processing pipeline, which processes captured images into extractable metrics. This image processing pipeline starts by filtering, cropping and sorting the pictures by tooth. This was followed by annotating the images using a Feature Pyramid Network, which was trained on 14 manually annotated images. These 14 images were annotated through an iterative process to reduce the annotation effort required, and upsampling of the training dataset was achieved by patchifying the images. Finally, using the annotations provided by the deep learning model and calculating the pixel size (22.4 μm), metrics were extracted from this dataset to analyse the spatiotemporal evolution of gear pitting during the performed test. This analysis concluded that for the generated pitting dataset, the evolution of the total pitted area could be divided into four zones, each relating to a different stage of gear health. The first zone was when the gear was healthy, followed by progressive macropits appearing at the 40-hour mark. This was followed by non-progressive macropits due to wear debris being caught between the meshing gears. Finally, spall pitting occurred due to the coalescing of pits. Hereafter, the distribution of the pitted area was analysed, and it was found that this distribution remains consistent as the test progresses and between different teeth, where the most significant difference is found for the larger pits. Similar trends were found for the amount of pits and the summed height and width. In addition, a relationship between the height and width of the pits was found, i.e., a height-to-width ratio of 77%. Finally, the position of the pits was investigated, and it was found that at the early stages, the pits appeared equally distributed around the same horizontal region of the gear flank. However, due to non-progressive macropitting resulting from wear debris, pits are found higher up the tooth flank at the later stages of the test.

Furthermore, this paper discussed the evaluation of the accuracy of this vision solution and found that despite challenges, such as underestimating pit boundaries and occasional failures, such as detecting small pits or wrong annotations due to temporary effects, the Intersection over Union (IoU) metric, ranging from 0.50 to 0.60 (median 0.56), was deemed sufficiently reliable for monitoring trends in pitting evolution.

For future research, additional tests should be performed to see if generalisable trends in the initiation and evolution of pitting can be observed. These trends could be helpful during the development phase of new systems to increase their lifetime. Additionally, these insights are attractive for modelling pitting as it would allow for more accurate models to be developed. Finally, generating additional data and performing tests on other surface wear mechanisms, such as abrasion and adhesion, could create a more generalized deep-learning model that could provide automated fault detection over various gearboxes. This would help the industry move towards more accurate and faster predictive monitoring of gearboxes, as the presented vision-based condition monitoring system can provide more detailed information on gear wear compared to the currently utilized techniques such as oil monitoring, vibration monitoring and acoustic emissions. An in-depth comparison between the visual monitoring solution and existing monitoring techniques is considered beyond the scope of this paper. However, synchronized measurements of vibrations, temperature, torque, and images of the worn gears were captured during the tests described in this work and are planned to be analyzed in a future study for a more detailed comparison.

**Acknowledgements** This research was supported by Flanders Make, the strategic research centre for the manufacturing industry, and explored during the SBO (Strategic Basic Research) project for QED (Quantified Evolution of Degradation of Gears, NO.: 2020-1138). Additionally, this author would like to thank the other partners in this project, as this work results from collaboration between all partners involved.

**Funding** The author acknowledges the funding provided by Flanders Make within the framework of the QED (Quantified Evolution of Degradation of Gears) SBO project (Strategic Basic Research), NO.: 2020-1138.

**Author Contribution** Djordy Van Maele, Jean Carlos Poletto, Dieter Fauconnier and Patrick De Baets contributed towards the ideation of the vision-based approach for monitoring gear health. Dieter Fauconnier and Patrick De Baets supervised the research. Djordy Van Maele and Jean Carlos Poletto performed the design of experiments, testing and collection of data. The conceptualization and development of the image processing pipeline by Djordy Van Maele, Jean Carlos Polleto, Roeland De Geest and Wenzhi Liao. Djordy Van Maele and Jean Carlos Polleto annotated the training dataset, while Roeland De Geest and Wenzhi Liao handled the deep-learning-related tasks. Djordy Van Maele, Jean Carlos Poletto, Dieter Fauconnier and Patrick De Baets dealt with the analysis of the results. Djordy Van Maele wrote the manuscript, and all authors contributed to reviewing and approving the final manuscript.

**Conflict of interest** D. Van Maele, J.C. Poletto, R. De Geest, W. Liao, N.F. Ferreira, P.D. Neis, D. Fauconnier and P. De Baets declare that they have no competing interests.



## References

1. Tchakoua P, Wamkeue R, Ouhrouche M, Hasnaoui F, Tameghe TAT, Ekemb G (2014) Wind turbine condition monitoring: state-of-the-art review, new trends, and future challenges. Energies 7:2595–2630. https://doi.org/10.3390/en7042595
2. Carroll J, McDonald A, Barrera O, Mcmillan D, Bakhshi R (2015) Offshore wind turbine sub-assembly failure rates through time.
3. Walford CA (2006) Wind Turbine Reliability: Understanding and Minimizing Wind Turbine Operation and Maintenance Costs. Tech Rep Sand2006-1100 Sandia Natl Lab 04:2024–2004. https://doi.org/10.2172/882048 (https://www.osti.gov/servlets/purl/882048)
4. Williams R, Zhao F (2023) Global offshore wind report 2023. Technical report, global wind energy council (GWEC) (August. https://gwec.net/wp-content/uploads/2023/08/GWEC-Global-Offshore-Wind-Report-2023.pdf. Accessed 2024-04-04
5. Sun J, Wang L, Li J, Li F, Li J, Lu H (2021) Online oil debris monitoring of rotating machinery: a detailed review of more than three decades. Mech Syst Signal Process 149:107341 (https://doi.org/10.1016/j.ymssp.2020.107341 . Accessed 2023-07-10)
6. Cao W, Zhang H, Wang N, Wang HW, Peng ZX (2019) The gearbox wears state monitoring and evaluation based on on-line wear debris features. Wear 426–427:1719–1728. https://doi.org/10.1016/j.wear.2018.12.068 (Accessed 2023-07-11)
7. Feng K, Ji JC, Ni Q, Beer M (2023) A review of vibration-based gear wear monitoring and prediction techniques. Mech Syst Signal Process 182:109605. https://doi.org/10.1016/j.ymssp.2022.109605 (Accessed 2023-04-05)
8. Feng K, Ji JC, Ni Q, Li Y, Mao W, Liu L (2023) A novel vibration-based prognostic scheme for gear health management in surface wear progression of the intelligent manufacturing system. Wear. https://doi.org/10.1016/j.wear.2023.204697 (Accessed 2023-04-03)
9. Zhang Y, Lu W, Chu F (2017) Planet gear fault localization for wind turbine gearbox using acoustic emission signals. Renew Energy 109:449–460. https://doi.org/10.1016/j.renene.2017.03.035 (Accessed 2024-04-05)
10. Gu D, Kim J, An Y, Choi B (2011) Detection of faults in gearboxes using acoustic emission signal. J Mech Sci Technol 25(5):1279–1286. https://doi.org/10.1007/s12206-011-0231-4 (Accessed 2024-04-05)
11. Kane PV, Andhare AB (2020) Critical evaluation and comparison of psychoacoustics, acoustics and vibration features for gear fault correlation and classification. Measurement 154:107495. https://doi.org/10.1016/j.measurement.2020.107495 (Accessed 2024-04-05)
12. Zhang R, Gu F, Mansaf H, Wang T, Ball AD (2017) Gear wear monitoring by modulation signal bispectrum based on motor current signal analysis. Mech Syst Signal Process 94:202–213. https://doi.org/10.1016/j.ymssp.2017.02.037 (Accessed 2023-07-23)
13. ISO (2021) ISO10825: gears—wear and damage to gear teeth—terminology.
14. Chen T, Zhu C, Chen J, Liu H (2024) A review on gear scuffing studies. Theor Exp Des 196:109741. https://doi.org/10.1016/j.triboint.2024.109741
15. Castro J, Seabra J (1998) Scuffing and lubricant film breakdown in FZG gears part II. New PV scuffing criteria, lubricant and temperature dependent. Wear 215(1-2):114–122. https://doi.org/10.1016/S0043-1648(97)00191-9
16. Castro J, Seabra J (1998) Scuffing and lubricant film breakdown in FZG gears part I. Analytical and experimental approach. Wear 215(1-2):104–113. https://doi.org/10.1016/S0043-1648(97)00277-9
17. Abdelbary A, Chang L (2023) 7—lubrication and surface engineering. In: Abdelbary A, Chang L (eds) Principles of engineering Tribology. Academic Press, pp 295–343 https://doi.org/10.1016/B978-0-323-99115-5.00012-8 (https://www.sciencedirect.com/science/article/pii/B9780323991155000128)
18. Xi D, Qin Y, Wang Y (2020) Vision measurement of gear pitting under different scenes by deep mask r-cnn. Sensors. https://doi.org/10.3390/s20154298
19. Xi D, Qin Y, Wang S (2023) Ydrsnet: an integrated yolov5-deeplabv3 + real-time segmentation network for gear pitting mea-

surement. J Intell Manuf 34:1585–1599. https://doi.org/10.1007/s10845-021-01876-y
20. Wang Z, Qin Y, Chen W (2021) Vision measurement of gear pitting based on dcgan and u-net. J Mech Sci Technol 35:2771–2779. https://doi.org/10.1007/s12206-021-0601-5
21. Miltenović A, Rakonjac I, Oarcea A, Perić M, Rangelov D (2022) Detection and monitoring of pitting progression on gear tooth flank using deep learning. Appl Sci. https://doi.org/10.3390/app12115327
22. Marian M, Bartz M, Wartzack S, Rosenkranz A (2020) Non-dimensional groups, film thickness equations and correction factors for elastohydrodynamic lubrication: a review. Lubricants. https://doi.org/10.3390/lubricants8100095
23. Höhn BR, Oster PTT, Hergesell M (2006) Short test procedure to investigate the lubricant influence on the pitting carrying capacity of gears. In:
24. De K, Masilamani V (2013) Image sharpness measure for blurred images in frequency domain. Procedia Eng 64:149–158. https://doi.org/10.1016/j.proeng.2013.09.086
25. Lin T, Dollár P, Girshick RB, He K, Hariharan B, Belongie SJ (2016) Feature pyramid networks for object detection (1612.03144)
26. lakubovskii P (2024) Segmentation model for python, based on Pytorch. https://smp.readthedocs.io/en/latest/models.html#fpn
27. Siegel C, Daily J, Vishnu A (2016) Adaptive neuron apoptosis for accelerating deep learning on large scale systems
28. Liao W, Geest RD (2024) Active learning for gear defect detection in gearboxes. Phm Soc Eur Conf 8:10. https://doi.org/10.36001/phme.2024.v8i1.4050
29. CVAT.ai: computer vision annotation tool (CVAT). https://github.com/opencv/cvat
30. Poletto JC, Fernandes CMCG, Barros LY, Neis PD, Pondicherry K, Fauconnier D, Seabra JHO, De Baets P, Ferreira NF (2023) Identification of gear wear damage using topography analysis. Wear. https://doi.org/10.1016/j.wear.2023.204837
31. Cardoso NFR, Martins RC, Seabra JHO, Igartua A, Rodríguez JC, Luther R (2009) Micropitting performance of nitrided steel gears lubricated with mineral and ester oils. Tribol Int 42(1):77–87. https://doi.org/10.1016/j.triboint.2008.05.010
32. Brandão JA, Seabra JHO, Castro J (2010) Surface initiated tooth flank damage. Part II: prediction of micropitting initiation and mass loss. Wear 268(1-2):13–22. https://doi.org/10.1016/j.wear.2009.07.020