Active learning for gear defect detection in gearboxes

Wenzhi Liao[1,2], Roeland De Geest[1], Djordy Van Maele[3], Jean Carlos Poletto[3], Laveen Prabhu Selvaraj[4], Ted Ooijevaar[1], Luk Geens[4]

[1] *Flanders Make, Oude Diestersebaan 133, 3920 Lommel, Belgium*
Wenzhi.liao@FlandersMake.be

[2] *IPI-TELIN, Ghent University, St-Pietersnieuwstraat 41, B-9000 Gent, Belgium*

[3] *Ghent University, Soete Laboratory, Technologiepark Zwijnaarde 46, 9052 Zwijnaarde, Belgium*

[4] *ZF Wind Power Antwerpen NV, Gerard Mercatorstraat 40, 3920 Lommel*

ABSTRACT

Condition monitoring of gears in gearboxes is crucial to ensure performance and minimizing downtime in many industrial applications including wind turbines and automotive. Monitoring techniques using indirect measurements (i.e. accelerometers, microphones, acoustic emission sensors and encoders, etc.) face challenges, including the defect interpretation and characterization. Vision-based gear condition monitoring, as a direct method to observe gear defects, has the capability to give a precise indication of the starting point of a potential surface failure, but suffers from the image annotations (to train a reliable vision model for automatic defect detection of gears). In this paper, we propose an active learning framework for vision-based condition monitoring, to reduce the human annotation effort by only labelling the most informative examples. In particular, we first train a deep learning model on limited training dataset (annotated randomly) to detect pitting defects. To select which samples have the highest priority to be annotated, we compute the model's uncertainty on all remaining unlabeled examples. Bayesian active learning by disagreement is exploited to estimate the uncertainty of the unlabeled samples. We select the samples with the highest values of uncertainty to be annotated first. Experimental results from defect detection of gears in gearboxes show that with less than 6 times image annotations, we can achieve similar performances.

1. INTRODUCTION

Detecting defects on gear surfaces is essential for maintaining the safety, performance, and longevity of machinery, while also ensuring quality control and minimizing downtime and costs, especially for gearboxes in high-power-density machines (e.g., wind turbines). Many approaches exploit indirect measurements acquired from accelerometers, microphones, acoustic emission sensors and encoders to monitor the damage evolution in gears (Surucu, Gadsden, & Yawney, 2023; Feng, Ji, Ni, & Beer, 2023). However, this indirect way of gear condition monitoring (e.g., vibration analysis) suffers from relative indicators and setting good thresholds to accurately track the gear damage (Surucu et al., 2023). Moreover, the indirect measurements cannot well characterize the defects (e.g., size, location, type) of the gears (Van Maele et al., 2023). Vision monitoring, which is a direct method to observe defects has the capability to give a precise indication of the starting point of a potential surface failure. Gear damage is often validated using visual inspection with borescopes or fibre scopes. However, such a system is used in some domains (mainly in wind turbines) as a periodic maintenance procedure but expensive equipment and permanent machine stop is needed (Coronado & Fischer, 2015). Recent advances in computer vision and machine learning have revolutionized industrial maintenance practices, allowing for the development of automated systems capable of visually inspecting and analyzing gear surfaces. Vision-based approaches utilize cameras and sensors to capture images or videos of gears during operation, enabling the extraction of meaningful visual features for condition assessment (Allam, Moussa, Tarry, & Veres, 2021; Qin, Xi, & Chen, 2023; Miltenović, Rakonjac, Oarcea, Perić, & Rangelov, 2022). This shift towards visual inspection not only facilitates continuous monitoring but also provides a more comprehensive understanding of gear health by capturing subtle surface details and anomalies. Massive image data can be acquired by high-speed cameras for visual condition monitoring of gears. Deep learning, particularly convolutional neural networks (Allam et al.,

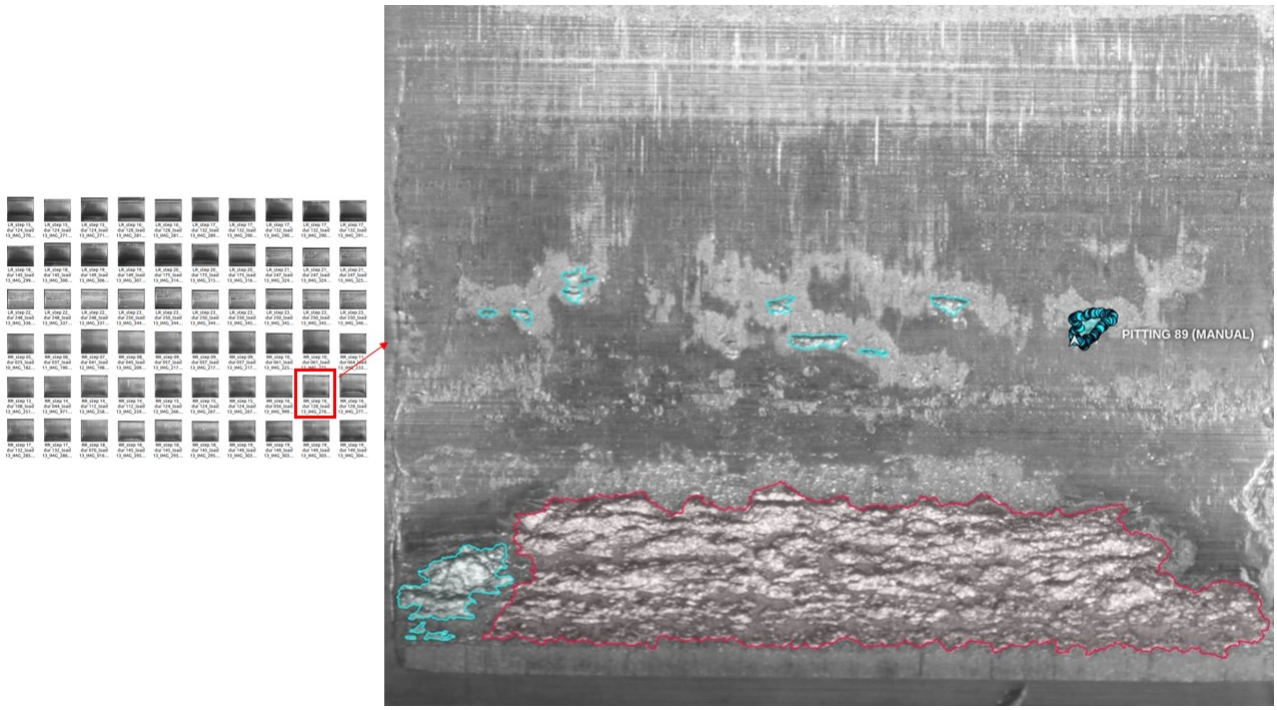


Figure 1. More than 20 minutes were taken by our experts to annotate defect of pitting in a single frame of image. Two types of defects were annotated for all images: micropitting (red), and pitting (cyan).

2021), has shown remarkable success in image-based tasks, making it well-suited for gear surface defect detection. However, to train a reliable vision deep learning model for automatic defect detection of gearboxes, a huge amount of image data typically needs to be annotated, which is expensive and time-consuming (Alzubaidi, Bai, Al-Sabaawi, & et al., 2023). For example, it takes more than 20 minutes to annotate all pitting defects in a single frame of image, as shown in Figure 1. Moreover, image datasets acquired during full lifetime degradation tests datasets contain many similar examples that bring no additional information to the diagnostic model. To overcome these problems, the active learning method was exploited to select the most informative indirect signals (e.g., vibration, supervisory control and data acquisition) for gearbox fault diagnosis (Chen et al., 2019) or wind turbine condition monitoring (Bao, Zhang, Hu, Feng, & Liu, 2023). Recent work on vision-based defect segmentation also showed that active learning framework can reduce data storage and labeling costs for imbalanced industrial datasets (Li et al., 2023).

To reduce the cost on manual annotation, this paper proposes an active learning framework to address the challenge of acquiring labeled data by iteratively selecting the most informative images for annotation. To the best of our knowledge, this paper is the first study to apply deep active learning for vision-based gear defect segmentation/detection in gearboxes. Specifically, a few images (i.e. around 20) were initially annotated to train a deep learning model for defect detection. To choose which gear images will be the first priority to be annotated, we then compute the model's uncertainty on all remaining unlabeled examples, where Bayesian active learning (Atighehchian et al., 2022) by disagreement is exploited to estimate the uncertainty of the unlabeled samples. The samples with the highest values of uncertainty will be chosen to be annotated first. We repeat the image annotations iteratively (e.g., top 10 images ranking according to the uncertainty will be annotated in each iteration) until we achieve a satisfactory performance.

The structure of this paper is as follows. Section 2 introduces the active learning framework. Section 3 details the experimental data collection and processing. The experimental results of defect detection on gear flanks are presented and discussed in Section 4. Finally, the conclusions of this paper are drawn in Section 5.

2. METHODOLOGY

2.1. Deep segmentation model

To monitor the damage evolution in gears, our solution first segments the damaged regions (defect) in the acquired images, then characterizes these damaged regions (change of size, shape, depth, etc.). A Python library with Neural Networks for Image Segmentation based on PyTorch (SMP) (Iakubovskii, 2019) is exploited for defect segmentation task

in this paper, as it is an open-source library built on top of PyTorch, specifically tailored for semantic segmentation tasks in computer vision. Semantic segmentation involves assigning a class label to each pixel in an image, thus dividing the image into distinct regions corresponding to different object classes. Semantic segmentation is additionally assigning each detected object a category and discriminates between objects of the same category. SMP includes an efficient and flexible implementation of Feature Pyramid Network (FPN) (Lin et al., 2017) for semantic segmentation tasks, combining low- and high-resolution features via a top-down pathway to enrich semantic features at all levels (multi-scale features). By leveraging multi-scale features and transfer learning, SMP-FPN enables accurate and robust segmentation of objects in images across various scales and contexts, fitting perfectly with the defect detection in the gears (defect area sizes changing).

The initial training dataset is very limited, since image annotation of these defect in the gears are challenging and time consuming. Therefore, we leverage pre-trained weights from models trained on large-scale image datasets such as ImageNet. The pre-trained weight of ResNet-18 (He, Zhang, Ren, & Sun, 2016), with a convolutional neural network that is 18 layers deep[1], is exploited in our segmentation model. The pre-trained model has been previously trained on more than a million images from the ImageNet database and contains the weights and biases that represent the features of whichever dataset it was trained on. These low-level learned features are often transferable to different data, including gears. For example, a model trained on a large dataset of natural objects (e.g., bird, fish images) will contain learned features like edges or textures that would be transferable defects in gears, which helps improve the performance of the segmentation model (especially with very small training sample size).

2.2. Active learning for image annotation

Even with a pre-trained model, the segmentation performances are still poor, especially for images mixed with two classes of "micropitting" and "pitting", as shown Figure 2, regions of micropitting were misclassified into pitting (poor performances in confusion matrix), while pitting defects were misclassified into background. An easy and simple solution to improve the performances is to add more annotated images into the training dataset. With a high-speed camera, we can acquire more than 60 image per second, around 30,000 images for 8 hours. However, image annotation is time consuming for our experts (an image shown in Figure 1 may take 20 minutes to annotate), even with advanced annotation tool CVAT [2]. Since it is infeasible for an expert to annotate all the acquired images, two challenges need to be solved: (1) which images should be first annotated? (2) how many images should be annotated for a reliable prediction?

Active learning aims to minimize the annotation effort required by selecting the most informative samples for annotation, i.e., the samples that would most increase the model accuracy. Active learning is a machine learning paradigm where a model iteratively queries the user or a human annotator for the labels of the most informative samples. This can lead to significant savings in time and resources compared to traditional approaches that rely on labeling large amounts of data upfront or passive learning from a fixed dataset.

Many active learning approaches have been proposed (Beluch, Genewein, Nürnberger, & Köhler, 2018; Kirsch, Amersfoort, & Gal, 2019; Wan et al., 2023), but some of these methods are either not scalable to large datasets or too slow to be used in a more realistic environment (e.g., in a production setup) (Atighehchian, Branchaud-Charron, & Lacoste, 2020). We exploit Bayesian Active Learning by Disagreement (BALD) (Atighehchian et al., 2020) in this paper to select the most informative samples for annotation. BALD leverages Bayesian modeling to estimate the uncertainty of a predictive model and selects samples where the model's predictions are most uncertain. In particular, BALD involves calculating the mutual information between the model's predictions and the model's parameters, given the observed data. Let D denote the labeled dataset, where $D = (\boldsymbol{x}_i, y_i)_{i=1}^{N}$ with inputs $\boldsymbol{x}_i$ and corresponding labels y_i. Let θ represent the model parameters, and $f_\theta(\boldsymbol{x})$ denote the predictive distribution of the model. The BALD acquisition function is defined as the mutual information between model parameters and potential labels of unlabeled data $\boldsymbol{x}$:

$$\begin{aligned} BALD(x) &= \boldsymbol{I}[y, f_\theta(\boldsymbol{x})] \\ &= \boldsymbol{H}[y] - \boldsymbol{E}_{p(f_\theta(\boldsymbol{x})|D)}[\boldsymbol{H}[y|f_\theta(\boldsymbol{x})]] \end{aligned} \tag{1}$$

Where:

- $\boldsymbol{I}[y, f_\theta(\boldsymbol{x})]$ is the mutual information between the label y and the model's prediction $f_\theta(\boldsymbol{x})$ for an unlabeled data point $\boldsymbol{x}$.
- $\boldsymbol{H}(y)$ is the entropy of the label distribution, measuring uncertainty in the label predictions.
- $\boldsymbol{E}_{p(f_\theta(\boldsymbol{x})|D)}$ is the expectation over the posterior distribution of the model given the current dataset D
- $\boldsymbol{H}[y, f_\theta(\boldsymbol{x})]$ is the conditional entropy of the label distribution given the model's prediction.

Intuitively, samples with higher BALD scores are those for which the model's predictions are most uncertain and thus are most informative for learning. By querying such uncertain samples, the model can learn more effectively with limited annotated data, leading to efficient data annotation for model training. In a normal image annotation task, our expert

[1] https://www.kaggle.com/datasets/pytorch/resnet18
[2] https://github.com/opencv/cvat

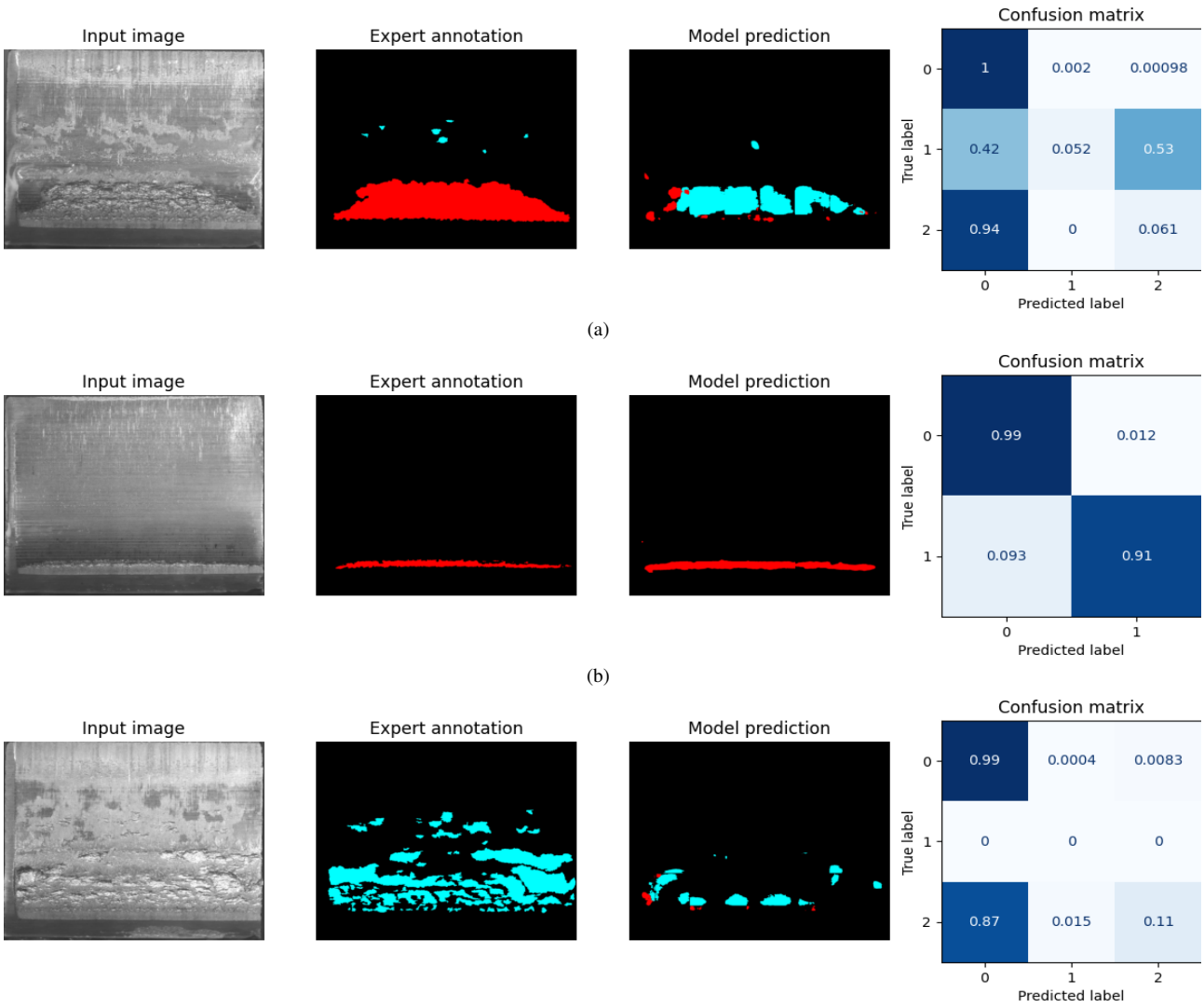


Figure 2. Performances on segmentation model trained on a small training dataset. The confusion matrices in the right column show the performances for three test images, 3 classes were defined by our experts in the images, with class label (color) 0: background (dark), 1: micropitting (red), 2: pitting (cyan).

annotators will start annotate images according to their order uploaded into a annotation tools (CVAT) or the project coordinator will assign a certain number of images randomly to each annotator. Compared to the active learning with random selection of samples for annotation, the uncertainty score of active learning with BALD tends to zero when reaching to 300 images in the first iteration, as shown in Figure 3. The active learning process using BALD is iterative. After annotating the selected samples and incorporating them into the training set, the model is retrained, and the process repeats. Over multiple iterations, the model becomes increasingly accurate, and the uncertainty decreases, leading to more confident predictions. The annotation loops will stop until the end-users satisfy with the performances, which can be evaluated by either through a matrix on validation dataset, or by manually interpretation on randomly selected images (if not enough validation reference images). Figure 3(b) shows that the uncertainty score of active learning with BALD tends to zero after 180 images in the second iteration, while active learning with random sample selection still needs to annotate all images to achieve this. By focusing on samples where the model's predictions are most uncertain, BALD enables efficient learning with limited annotated data.

Table 1. Acquired images and manual annotations.

No. Teeth	No. Annotated teeth	No. Images	No. Annotated Images	No. Annotated Polygons
54	18	1370	438	1036

Table 2. Data split (within 18 annotated teeth, 438 annotated images) for active learning.

Initial training dataset	Tooth 1 (23 images)
Validation dataset	Tooth 15 (31 images)
Test dataset	Tooth 5 (31 images)
Pool dataset	The other 15 teeth (353 images)

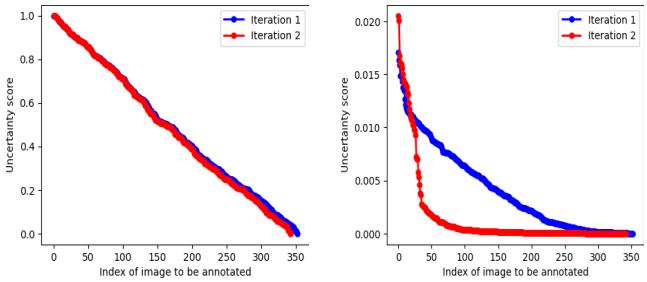


Figure 3. Uncertainties scores by active learning with: (a) random selecting samples for annotation, (b) BALD. Iteration 1 means 10 annotated images are added into the initial training dataset and retrain the model of active learning; iteration 2 means 20 images are iteratively annotated and added into the initial training dataset.

3. Experiments Data Collection and Processing

3.1. Data Collection

A dataset containing images of gears subjected to accelerated lifetime tests was provided by ZF Wind Power. While a brief description of the dataset acquisition is present here, the reader is referred to the work of (Boemher, 2019) for further details. The dataset consists of two accelerated lifetime tests, and was generated on a standard FZG[3] gear test rig at ZF Wind Power. At selected moments of the test, the equipment was stopped and images of both meshing gears were manually captured using a Canon EOS 500D camera. Figure 4 illustrates the gear degradation of a gear flank throughout the test. Two pairs of standard FZG C14 spur gears with 16 teeth (pinion) and 22 teeth (wheel) were tested on each accelerated lifetime test. In the first test, with total duration of $152h$, the test was stopped 31 times for acquiring the image of the gear flanks. Meanwhile, on the second test with total duration of $250h$, image acquisition was performed 23 times. A prior qualitative assessment determined that the wheel of the first test did not developed damage. Hence, the assessed dataset is composed of 54 teeth: pinion (16) of first test, plus pinion (16) and wheel (22) of the second test.

[3]Forschungsstelle für Zahnräder und Getriebebau, which denotes the Gear Research Center at the Technical University of Munich

3.2. Experimental Setup

The accelerated lifetime testing procedure was designed to generate micropitting and pitting wear on the visually monitored gear surfaces. As shown in Table 1, 54 teeth were used in experiments and a large amount of images was acquired by our camera. After filtering the unclear images (i.e., blurry, noisy, etc.) and pre-processing, we obtain 1370 images, of which 438 images were annotated by our experts, as shown in Table 2. The annotation effort varied according to the amount of defects in each image, taking approximately 60 hours to fully annotate the dataset (438 images), and in some cases up to 30 minutes were required to annotate a single frame.

Two sets of experiments are compared:

- **Fixed_SMP**: train the SMP-FPN models using Initial training dataset + Pool dataset, totally 16 teeth, 376 annotated images;
- **Active_SMP**: train the SMP-FPN model initially on Initial training dataset (Tooth 1, with 23 annotated images in total)

Then a number of annotated samples (i.e. 10 in each iteration) selected by active learning from the Pool dataset are iteratively added into the training dataset, and the model is retrained. Within active learning segmentation, we will compare different methods to select samples for first priority to be annotated, such as:

- **Active_SMP_Random**: select 10 images randomly in each iteration;
- **Active_SMP_BALD**: select 10 images by using BALD method in each iteration.

We set some parameters for model training as: batch size: 8, epochs: 100, learning rate: 0.0001. To reduce inherent randomness in the training of deep networks, each experiment runs five times for active learning.

For performance evaluations, we exploit the confusion matrix (Powers, 2011) to report the performance of a segmentation model on a single image. This confusion matrix helps in understanding where the segmentation model is making errors, whether it is under-segmenting or over-segmenting certain classes, or if there are misclassifications between classes. It is an essential tool for evaluating the effectiveness of segmentation algorithms. To evaluate the segmentation models on the whole test dataset, we exploit mean intersection over

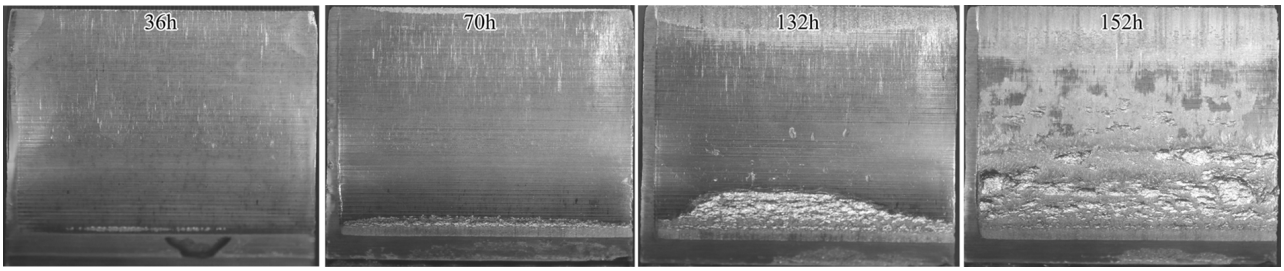


Figure 4. Example of images collected at selected moments of the test, showing the evolution on the gear degradation with the test duration.

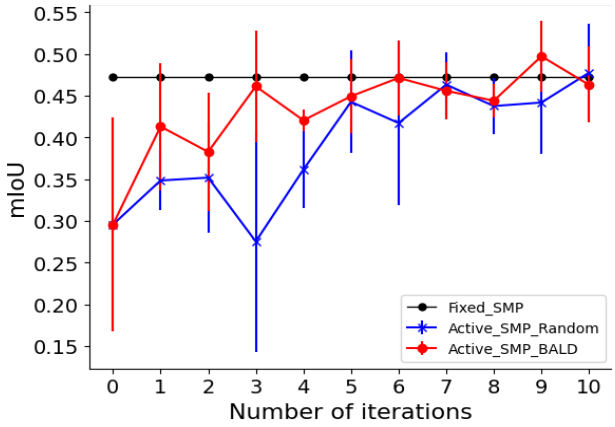


Figure 5. Mean and standard deviation mIoU of different segmentation methods. Note: the Fixed_SMP method is trained by using fixed number of 376 annotated images; iteration 0 means the model of active learning is trained on Initial training dataset (Tooth 1 with 23 images); 10 images will be selected in each iteration for annotation and then added into the training dataset, the model with active learning will be retrained (for example iteration 3 means 30 images are iteratively annotated and added into the Initial training dataset). We repeated the active learning segmentation experiments 5 times.

union (mIoU), which measures the overlap between the predicted segmentation and the ground truth segmentation for each class or object in the image (with Python implementation[4]). The value of the metric ranges from 0 to 1, higher value indicates better performance on segmentation.

4. Results and Discussions

Figure 5 compares the performances of segmentation models trained by Fixed_SMP and Active_SMP. The changes of the predicted segmentation maps by adding more annotated images into the training dataset can be found in Figure 6. We take several test images as examples and show the segmentation results and their confusion matrices by using Fixed_SMP and active SMP_BALD in Figure 7.

Based on Figure 5, we can find that with 53 annotated images (i.e. 3 iterations), active learning with BALD can achieve similar performances as Fixed_SMP (where more than 360 annotated images are used for training), which requires 6 times less annotated images for training, reducing more than 6 times the manual annotation effort. Moreover, Active_SMP_BALD performs better than Active_SMP_Random, especially for the first 5 iterations, when a small number of images are selected for annotations. This means that BALD can select the most informative images (out of a large dataset) for annotation when limited manpower is available for annotation. The model can learn more effectively with fewer BALD selected images, leading to efficient data annotation for model training. As more annotated images (more than 70 annotated images) are added into the training dataset, Active_SMP_Random converges to similar performances as the method of Fixed_SMP, indicating the redundancy in the image annotations. This is because images acquired during full lifetime degradation for multiple teeth contain many similar defects that bring no additional information for model training.

The segmentation models with active learning becomes more stable, as more annotated images added into the training dataset, as indicated by the changes of standard deviation in Figure 5. However, there are scenarios where increasing the training sample size (by adding more annotated images) might seemingly degrade segmentation performance, defects of “micropitting” appear in Figure 6-7 (Fixed_SMP for the third image) as more annotated images added. This may be due to overfitting, the model should generalize better with more training samples (that have similar distributions as the test images). One solution is to add more images that are representative of different classes into the pool dataset for active learning.

Compared to the human expert annotations, the segmentation results predicted by deep learning models (even for Fixed_SMP) need to be improved, regions of “micropitting” and “pitting” are misclassified into background, whereas some background regions are also misclassified into “micropitting”, as indicated by the confusion matrices in Figure 7. Challenges remain in predicting very tiny ”pitting” defects, as well as images mixed with big “micropitting” defects and

[4]https://lightning.ai/docs/torchmetrics/latest/segmentation/mean_iou.html

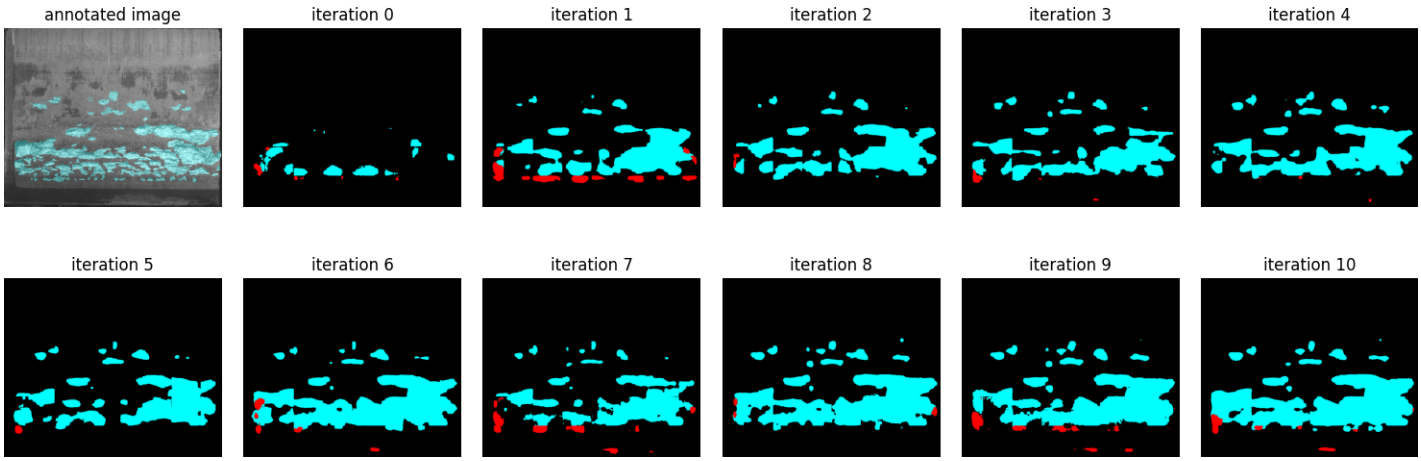
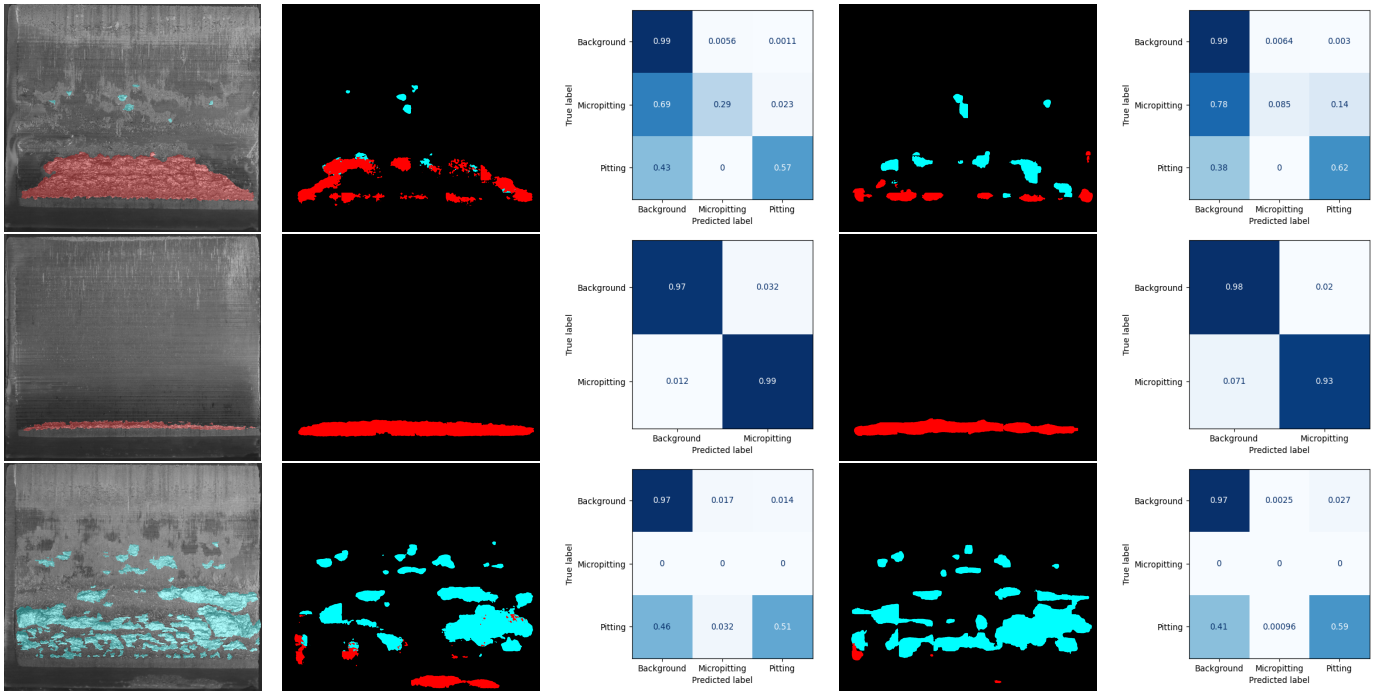


Figure 6. Prediction map changes as adding more annotated images (selected by BALD) into training dataset. 10 images will be annotated in each iteration and then added into the training dataset.



Expert annotations | Fixed_SMP | Fixed_SMP | active_SMP_BALD | active_SMP_BALD

Figure 7. Performances on segmentation by fixed training number VS. active learning. Each row has one test image, column 1 shows highlighted annotated images by experts, column 2 and 3 show predicted segmentation maps and confusion matrices by Fixed_SMP, column 4 and 5 are predicted segmentation maps and confusion matrices by Active_SMP_BALD with three iterations (53 training images).

tiny ”pitting” defects.

5. Conclusions

This paper focus on training a reliable deep learning segmentation model for defect detection in gears using less image annotations. In particular, Bayesian Active Learning by Disagreement (BALD) was exploited to select the most informative images for annotation iteratively until the satisfied performances were achieved. Experimental results show that with less than 6 times image annotations, we can achieve similar performances, leading to significant savings in time and resources compared to traditional approaches that rely on labeling large amounts of data upfront. However, gear surfaces exhibit a variety of defect types and patterns, and the successful identification of these defects requires a model capable of learning intricate features and subtle variations. The initial results in this paper can be extended by considering: (1) uncertainties from human annotations (annotations may be different by different human annotators in Figure 8), (2) imbalance in class distribution (some classes have more annotations than the other classes), (3) data augmentation to increase diversity in the training image dataset, (4) validation of the active learning methods on wider applications using some public datasets (e.g., ball screw drive surface defect dataset (Schlagenhauf & Landwehr, 2021)) for more comprehensive comparisons.

Acknowledgment

This research was supported by Flanders Make, the strategic research centre for the manufacturing industry, and more precisely the SBO (Strategic Basic Research) project for QED (Quantified Evolution of Degradation in gears, NO.: 2020-1138). The authors would like to thank ZF Wind Power NV for providing the gear image dataset that was used in this research.

References

Allam, A., Moussa, M., Tarry, C., & Veres, M. (2021). Detecting teeth defects on automotive gears using deep learning. *Sensors*, *21*(24).

Alzubaidi, L., Bai, J., Al-Sabaawi, A., & et al. (2023). A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. *Journal of Big Data*, *10*(46).

Atighehchian, P., Branchaud-Charron, F., Freyberg, J., Pardinas, R., Schell, L., & Pearse, G. (2022). *Baal, a bayesian active learning library.* `https://github.com/baal-org/baal/`.

Atighehchian, P., Branchaud-Charron, F., & Lacoste, A. (2020). *Bayesian active learning for production, a systematic study and a reusable library.*

Bao, C., Zhang, T., Hu, Z., Feng, W., & Liu, R. (2023). Wind turbine condition monitoring based on improved active learning strategy and knn algorithm. *IEEE Access*, *11*, 13545-13553.

Beluch, W., Genewein, T., Nürnberger, A., & Köhler, J. (2018). The power of ensembles for active learning in image classification. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 9368–9377).

Boemher, D. E. (2019, 4). Computer vision for gear alignment check and condition monitoring of wind turbine gearboxes.

Chen, J., Zhou, D., Guo, Z., Lin, J., Lyu, C., & Lu, . (2019). An active learning method based on uncertainty and complexity for gearbox fault diagnosis. *IEEE Access*, *7*, 9022-9031.

Coronado, D., & Fischer, K. (2015). Condition monitoring of wind turbines : State of the art , user experience and recommendations project report..

Feng, K., Ji, J. C., Ni, Q., & Beer, M. (2023). A review of vibration-based gear wear monitoring and prediction techniques. *Mech. Syst. Signal Process.*, *182*, 109605.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 ieee conference on computer vision and pattern recognition (cvpr)* (p. 770-778).

Iakubovskii, P. (2019). *Segmentation models pytorch.* https://github.com/qubvel/segmentation_models.pytorch. GitHub.

Kirsch, A., Amersfoort, J., & Gal, Y. (2019). *Batchbald: Efficient and diverse batch acquisition for deep bayesian active learning.*

Li, W., Li, B., Niu, S., Wang, Z., Liu, B., & Niu, T. (2023). Selecting informative data for defect segmentation from imbalanced datasets via active learning. *Advanced Engineering Informatics*, *56*, 101933.

Lin, T., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In *2017 ieee conference on computer vision and pattern recognition (cvpr)* (p. 936-944).

Miltenović, A., Rakonjac, I., Oarcea, A., Perić, M., & Rangelov, D. (2022). Detection and monitoring of pitting progression on gear tooth flank using deep learning. *Applied Sciences*, *12*(11).

Powers, D. (2011). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *ArXiv*, *abs/2010.16061*.

Qin, Y., Xi, D., & Chen, W. (2023). Gear pitting measurement by multi-scale splicing attention u-net. *Chinese Journal of Mechanical Engineering*, *36*(50).

Schlagenhauf, T., & Landwehr, M. (2021). Industrial machine tool component surface defect dataset. *Data in Brief* , *39*, 107643.

Surucu, O., Gadsden, S. A., & Yawney, J. (2023). Condition monitoring using machine learning: A review of the-

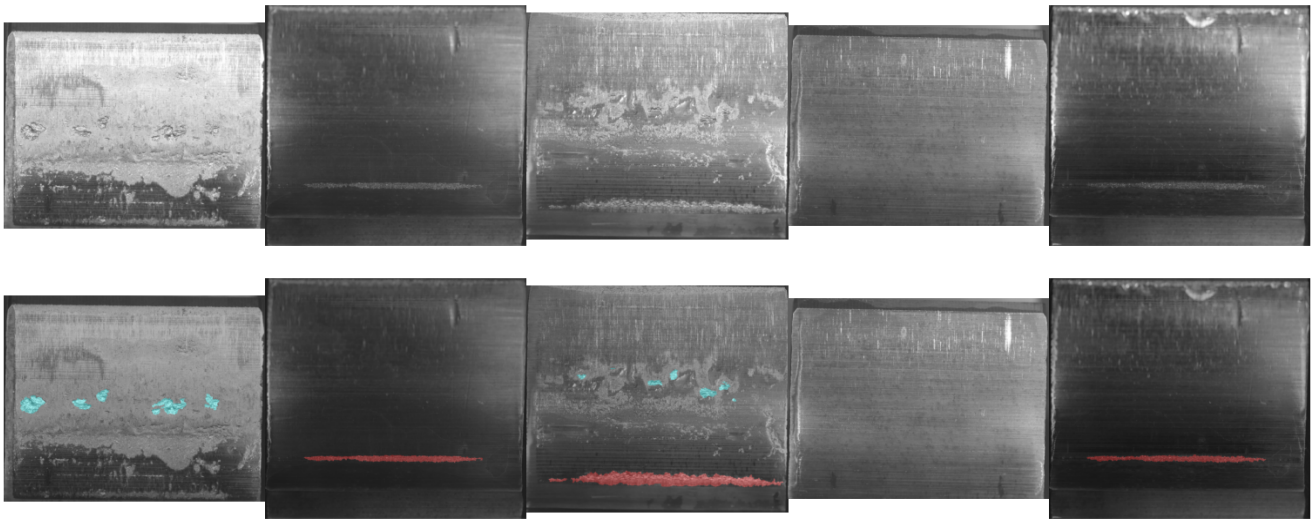


Figure 8. Uncertainties in image annotation. The first top 5 images (row 1) selected by active learning with BALD to be annotated, and their annotations (annotated by our experts in row 2).

ory, applications, and recent advances. *Expert Systems with Applications*, *221*, 119738.

Van Maele, D., Poletto, J. C., Neis, P., Ferreira, N., Fauconnier, D., & De Baets, P. (2023). Online vision-assisted condition monitoring of gearboxes. In *8th euro. conf. and exhibition on lubrication, maintenance and tribotech (lubmat 2023)*.

Wan, T., Xu, K., Yu, T., Wang, X., Feng, D., Ding, B., & Wang, H. (2023). A survey of deep active learning for foundation models. *Intelligent Computing*, *2*, 0058.

BIOGRAPHIES

Wenzhi Liao received the Ph.D. degree in Engineering from the South China University of Technology, Guangzhou, China, in 2012, and the Ph.D. degree in computer science engineering from Ghent University, Ghent, Belgium, in 2012. From 2012 to 2019, he has been working first as a Postdoc at Ghent University and then as a Postdoctoral Research Fellow for Research Foundation Flander (FWO). From February 2020 to January 2022, he had worked in VITO (Mol, Belgium) as a Data Scientist. Since February 2022, he works in Flanders Make, focusing on smart vision for Industry 4.0. His current research interests include Image Processing and Interpretation, Pattern Recognition, AI and Computer Vision. He is also highly experienced in Machine Learning, Large-scale problems and Remote Sensing. Dr. Liao was a recipient of the Best Paper Challenge Awards in both the 2013 IEEE GRSS Data Fusion Contest and the 2014 IEEE GRSS Data Fusion Contest. He serves as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS).

Roeland De Geest is a researcher at Flanders Make with 6+ years of experience in industry-driven research in the field of computer vision and machine learning. He obtained a Ph.D. degree in Electrical Engineering from the KU Leuven in 2019.

Djordy Van Maele Received his M. Sc. in electromechanical engineering technologies from Ghent University, Gent, Belgium, in 2021. In 2021 he started working on his doctoral degree in electromechanical engineering, in the field of tribology, at Ghent University until current date.

Jean Carlos Poletto obtained his M.Sc. degree in Mechanical Engineering from the Federal University of Rio Grande do Sul (UFRGS), Brazil, in 2018. He has been developing research on experimental tribology, with active contributions to the field since 2015. Currently, he is working on a joint PhD program between UFRGS, Brazil, and Ghent University, Belgium.

Ted Ooijevaar is Senior Technology Domain Leader at Flanders Make with 8+ years of experience in industry-driven research and development for aerospace, machine and vehicle applications. Leads a research team focused on sensing (sensor fusion, virtual sensing), data analytics, modeling (physics, AI and physics augmented AI), monitoring (anomaly detection, diagnostics and prognostics) and health management technologies for machines and vehicles. Prior to this, he performed industry-driven research in the field of condition monitoring and data analytics as (senior) research engineer. He holds a Ph.D. degree in Mechanical Engineering from the University of Twente in the Netherlands in 2014. He also gained experience as a visiting Ph.D. researcher at the Nondestructive Evaluation Sciences Branch at the NASA Langley Research Center in the USA.

Laveen Prabhu Selvaraj is Senior Digital Solutions Engineer at ZF Wind Power Antwerpen NV from 2019. He is working on Condition Monitoring Systems (CMS) for Wind turbine gearbox, also working on new technology, sensor and innovation in WTG CMS systems. He was working on development of BIO sensors and characterization of Deep UV LEDS using electro luminescence spectroscopy (µEL) during his work as Scientific Researcher at Chemnitz University of Technology, Chemnitz, Germany for 3 years from 2016. He obtained his M.Sc. degree in Micro and Nano system from Chemnitz University of Technology, Chemnitz, Germany.

Luk Geens is a Senior Technology Engineer at ZF Wind Power Technology Antwerp (Belgium) with 20+ years of experience in wind turbine industry.