AI for NDE 4.0 – Recent use cases

Tuomas Koskinen[a], Topias Tyystjärvi, Oskar Siljama, Iikka Virkkunen

Trueflaw Oy, Tillinmäentie 3, tila A113, FIN-02330 Espoo, Finland
[a]Email : tuomas.koskinen@trueflaw.com

Abstract

The use of machine learning in non-destructive evaluation (NDE) is a growing trend in the industry and a necessary development towards NDE 4.0. Beside academia, there are numerous case examples where machine learning is in use for actual inspections already. The main benefit of machine learning powered NDE is its reliability and repeatability to find flaws from the data. However, as the fundamental is in image recognition, machine learning can facilitate the inspection in general beyond just finding defects. These can be recognising the welds, image quality indicators and other features which the inspectors usually have to identify by themselves. In addition to reliability increase, automating these repeatable tasks increase the speed of data analysis considerably, saving inspector's time where it is most valuable. In this paper we review use cases where machine learning has been used in NDE and how these approaches benefit the end customer.

1.0 Introduction

Traditionally for mechanized NDE, inspector goes through all the data manually. This is a laborious task and usually takes a lot more time than the actual recording of the data. Naturally this data amount has been already limited by scanning only areas which are structurally most significant or are known to be under considerable loads or corrosive environments. However, there is still considerable amount of data to be evaluated by the inspector and the flaws are typically rare. Thus, the majority of the inspector's time is spent viewing data that contain no flaws. Furthermore, as inspectors' time is spent in data evaluation the final report is often limited to the bare essentials regarding data evaluation; pass or fail.

Machine learning (ML) changes this considerably. The ML model can go through the data and highlight the areas of interest. This leaves the inspector to only look and evaluate the possible flaw indications, skipping the looking of data which contains no indications. Moreover, ML can be used as an extra inspector where the procedure requires one or more inspectors to go through the data.

Previous research cases have covered ultrasonic data in welds [1] and noisy conditions [2]. Machine learning models have also been successfully in more challenging welds such as austenitic welds [3] and dissimilar metal welds [4]. The field progresses forward with modern ML models as [5] used modern U-net [6] for feature detection in radiographic images and [7] compared the effectiveness of other modern models such as ResNet and MobileNet with ultrasonic data.

In ultrasound NDE, ML has been used beside flaw detection in denoising of the signal to achieve higher image quality with autoencoders [8, 9, 10]. Naturally, this autoencoder approach has been used for increasing the efficiency of flaw detection with a ML model [11] and artefact detection [12].

Modern ML approaches have been proven to work in NDE for feature and flaw detection. Furthermore, they have proven to be versatile enough to process the recorded data even further. However, while the ML models are ready and capable for field use, the mindset needs to change in order to fully achieve all the benefits from NDE 4.0. As machine learning has enabled a great reliability and efficiency increase in presentation needs to be taken even further. Traditionally, inspections have provided a report with a pass or fail evaluation where the data might not have been recorded. In [13] visions NDE 4.0 to be driven more toward data-oriented manner, where inspection and sensory data could be readily implemented to process monitoring. However, the transition needs to be gradual or the sudden requirement of cognitive demands and complete paradigm change might cause lack of acceptance or insufficient use [14].

In this paper we demonstrate the capability of modern machine learning models for ultrasonic inspection with two different model architectures, VGG16 [15] and U-net. Furthermore, we demonstrate the data presentation and reporting and how it evolves from traditional report to more interactive way of viewing the inspection results.

2.0 Use Cases

For the use cases we go through four ML solutions for NDE. First the metro axle inspection with ultrasound, second control-rod drive mechanism inspection with time-of-flight diffraction, third dissimilar metal weld inspection with phased array ultrasound, and finally a composite inspection with ultrasound.

2.1Metro axle inspection

For following case studies, the earliest model was taken into use in 2019, when first metro axle inspection was made with the aid of a machine learning model. The inspection is made with conventional ultrasound but with multiple probes. The data in new axles are simple, but as the axles wear during use, indications start to show from dents from rocks. Varying geometry makes the data interpretation challenging even for an experienced inspector. Due to these aspects, the use of simple amplitude threshold is infeasible, since the inspector needs to evaluate the indication and distinguish the signal from other indications.

A machine learning model was taught to detect cracks from the ultrasonic data. The model was based on a VGG-16 image recognition model. Thermal fatigue cracks were made into an axle specimen to generate initial data for machine learning. The data was augmented with virtual flaws to generate ample data and extra thermal fatigue cracks were manufactured for testing and validation of the model. The working model managed to reduce the inspection time considerably from multiple hours to an hour. While the increase in effectiveness was considerable, the reliability of the model was also significant.

The approach enabled data recording as previously, but in the data analysis the inspector was aided by the machine learning model and would only look at sections from the data which the model highlighted as seen in Figure 1. While the model did not point the exact location of the flaw it screened area of the data which the inspector would the evaluate and determine if the area contained a flaw. The final output of the model is a traditional PDF file, highly similar to normal inspection report. This output and functionality is industry 3.0 as there is no interaction in the report or automation for the data to be uploaded to a network for further storage and use.

Figure 1: Output for the axle inspection. Indications highlighted in red from the calibration axle.

2.2 Control-rod drive mechanism inspection (CRDM)

CRDM inspections in nuclear power plants are done with time-of-flight diffraction (TOFD) technique. Due to the geometry of the drive mechanism, there are a lot of indications that are not actually flaws. The data evaluation is challenging as the diffraction tip signal is difficult to detect when scrolling through a considerable amount of TOFD data. In addition, there can be noise from poor contact or similar deviations.

The approach was similar as with the metro axle case, since the model was based on the VGG-16 approach. The model would screen the data and highlight areas for the inspector to look at more carefully. As the inspector is no longer required to scour all the data as in Figure 2, the risk of missing a flaw by a mistake is reduced. This is particularly helpful since the data is required to be evaluated by multiple different inspectors, thus the time from one evaluation can be decreased considerably.

This model was put to a field trial, where number of qualification flaws were evaluated by the model and the performance was compared to a human inspector evaluating the same data.

The model was able to detect the same qualification flaws as the human inspector. No qualification flaws were missed by the human inspector or the machine learning model. However, there was a significant difference in the evaluation time between the inspector and the inspector aided by the machine learning model.

The output report seen in Figure 3 was an HTML document, where the results were highlighted for the inspector and the inspector could scroll through the vicinity of the indication to verify the result. Slight improvement toward NDE 4.0, but still the automation and system integration aspects could still be enhanced to achieve true NDE 4.0 performance

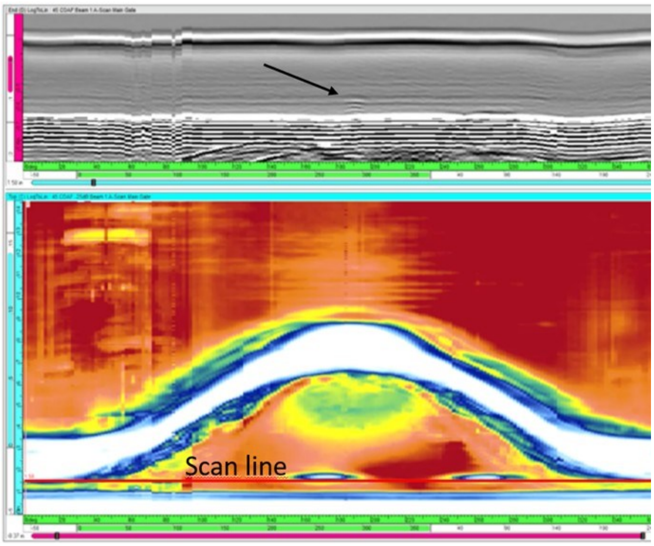

Figure 2: Traditional way of inspecting TOFD data

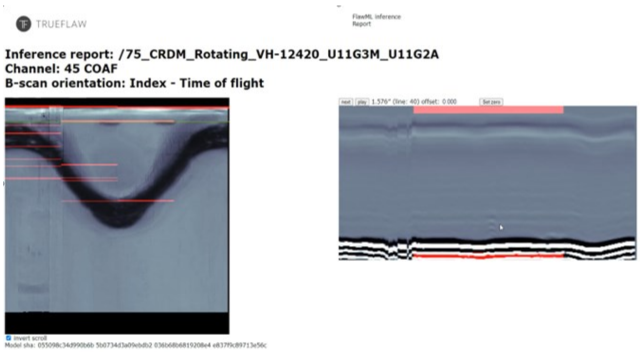

Figure 3: Indications highlighted by the ML model

2.3 Dissimilar metal weld inspection

Dissimilar metal weld (DMW) inspections are notoriously difficult inspections for ultrasound due to high noise and attenuation from the material structure. Moreover, these inspections are usually made in safety critical areas with high demand on reliability and low tolerance on false calls. Mostly these inspections are done with phased array ultrasound to gather enough information for flaw detection. As the data amount increases, so does the evaluation time. To reduce and ease the evaluation, different image and data processing methods are commonly used to help the inspector. One of these is merging the different angles obtained from phased array data to form a single merged ultrasonic image for the inspector to view. However, as the merge methods tend to leave artifacts and there is a possibility to lose data in the process, inspector usually needs to view the data completely.

For this case a U-net based semantic-segmentation model was used to detect flaws from the ultrasonic data and the output from the model can be seen in 4. In addition to detecting the flaws, the model could indicate the location of the flaw. By knowing the flaw location data analysis could be facilitated in such a way that an indication list would be presented to the inspector. This indication list would then show the actual indication in the data with pre-processed ultrasonic data. Then the inspector could readily measure the indication and move to the next one. As in the metro axle and CRDM inspection cases, the inspector would only evaluate the indications shown by the model, increasing the data evaluation speed considerably. Furthermore, the approach was more integrable with other systems as the result could be implemented to a regular inspection software as well if further analysis would be required.

2.4 Composite inspection

For the composite inspection the model was based on the VGG-16 approach. The target was to detect pores formed in the composite manufacturing. While ultrasound is ideal to detect pores and laminar defects from composite structures, the manual labour of the data analysis makes the pore detection and monitoring infeasible for human inspectors. Figure 5 demonstrates how the ML model has detected and highlighted the pores in the specimen.

As seen in Figure 5 there are numerous indications, thus it would have been infeasible for a human inspector to mark and track all the indications.

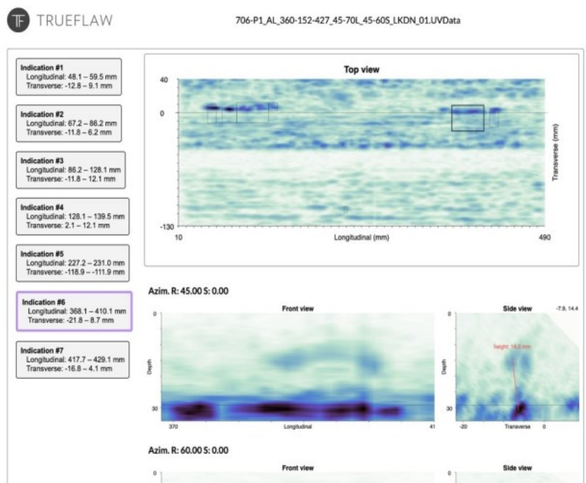

Figure 4: Result for DMW inspection from the ML model. Indication list on the left, which highlights the findings from the data. The height of the flaw can be measured easily by the inspector.

Once these indications are known they can be tracked over production batches for quality control purposes. Moreover, if pores pack as denser cluster they can be viewed as a defect. For the pore amount to be used in production statistics or monitoring, the data is possible to be integrated and implemented automatically to other production software to easily monitor the production parameters.

3.0 Discussion

The most recognizable benefit of these machine learning powered inspections is the efficiency. For most cases there is no need for the inspector to look at data that contains no indication about flaws or flaws, thus the result is promptly available.

Beside the considerable efficiency increase, the other significant benefit is the repeatability. Once the model has passed the qualification, the performance stays the same for future inspections. The performance stays predictable and there is no variation between evaluations. When compared to a human inspector, the performance of the evaluation is highly related to the inspector's skill but also how mentally fit the inspector is to do the evaluation endeavour.

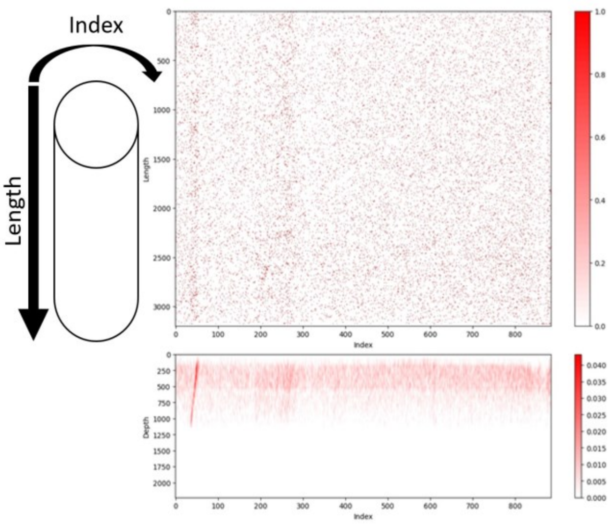

Figure 5: Composite inspection with a machine learning model. Areas marked in red are tiny pores from the specimen.

The repeatability is also highly linked to the overall reliability increase of the inspection. This reliability is measurable trough Probability of Detection (POD), which is well standardized practice in NDE [16]. POD elaborates the flaw size, which can be found the most feasibly, determining the critical flaw size. Auspiciously, missing of smaller flaws, is not significant comparing to a miss of a larger, critical sized one. Thus, the qualification effort needs to be set so, that the critical flaw size is found with high confidence. Through this high confidence, the inspector is always given the chance to evaluate the flaws within the scope of critical flaw sizes.

Once the inspector is given the chance to evaluate all critical sized flaws and possibly variety of smaller flaws the inspection effort is still highly controlled by the inspector. This enables the inspectors to use their time more efficiently by evaluating actual indications and not looking for them in the data, as they could miss some of the indications.

Finally, the ML opens completely new possibility for NDE by enabling larger variety of monitoring of the product. Features which are previously too laborious to monitor can now be implemented with ease with the aid of ML. This can either be implementing this monitoring to ongoing inspections or either inspecting the product with the goal of calculating porosity or other features. Furthermore, these approaches can be further integrated by automatically outputting the results in machine readable format for other systems to utilize and take advantage of.

4.0 Conclusion

Machine learning models have come to the NDE field and number of them are already in use. The models and approaches have matured to a point, where the machine learning models have passed the field trials in highly challenging data evaluations. Thus, the approaches are ready to increase the reliability of even the most safety critical industries.

The main benefit of utilising machine learning in NDE is the significant increase in efficiency, but also in reliability. When the inspector is able to focus to look at the data which is probable to have a flaw, the inspector's time is used more sensible way while still maintaining the control with the inspector.

Machine learning powered inspections also enable completely different approach to inspections in general. Once the data analysis can be automated, it enables monitoring of manufacturing in completely different scale than before, completely unreachable by manual data analysis.

Acknowledgements

The metro axle inspection was done in collaboration with DEKRA, CRDM and DMW inspection with EPRI and the composite inspection with Valmet. Their contribution is greatly appreciated.

References

[1] Nauman Munir, Hak-Joon Kim, Sung-Jin Song, and Sung-Sik Kang. Investigation of deep neural network with drop out for ultrasonic flaw classification in weldments. *Journal of Mechanical Science and Technology*, 32(7):3073–3080, 2018.

[2] N Munir, HJ Kim, J Park, SJ Song, and SS Kang. Convolutional neural network for ultrasonic weldment flaw classification in noisy conditions. *Ultrasonics*, 2018.

[3] Oskar Siljama, Tuomas Koskinen, Oskari Jessen-Juhler, and Iikka Virkkunen. Automated flaw detection in multi-channel phased array ultrasonic data using machine learning. *Journal of Nondestructive Evaluation*, 40(67), 2021.

[4] Tuomas Koskinen, Iikka Virkkunen, Oskar Siljama, Oskari Jessen-Juhler, and Jari Rinta-Aho. The effect of different flaw data to machine learning powered ultrasonic inspection. *Journal of Nondestructive Evaluation*, 40(24), 2021.

[5] Topias Tyystj¨arvi, Iikka Virkkunen, Peter Fridolf, Anders Rosell, and Zuheir Barsoum. Automated defect detection in digital radiography of aerospace welds using deep learning. *Welding in the World*, 66, 2022.

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[7] Luka Posilovi´c, Duje Medak, Fran Milkovi´c, Marko Subaˇsi´c, Marko Budimir, and Sven Lonˇcari´c. Deep learning-based anomaly detection from ultrasonic images. *Ultrasonics*, 124:106737, 2022.

[8] Dimitris Perdios, Adrien Besson, Marcel Arditi, and Jean-Philippe Thiran. A deep learning approach to ultrasound image recovery. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, 2017.

[9] Bo Li, Kele Xu, Dawei Feng, Haibo Mi, Huaimin Wang, and Jian Zhu. Denoising convolutional autoencoder based b-mode ultrasound tongue image feature extraction. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7130–7134, 2019.

[10] Fei Gao, Bing Li, Lei Chen, Xiang Wei, Zhongyu Shang, and Chen He. Ultrasonic signal denoising based on autoencoder. *Review of Scientific Instruments*, 91(4):045104, 2020.

[11] Nauman Munir, Jinhyun Park, Hak-Joon Kim, Sung-Jin Song, and SungSik Kang. Performance enhancement of convolutional neural network for ultrasonic flaw classification by adopting autoencoder. *NDT E International*, 111:102218, 2020.

[12] Sergio Cantero-Chinchilla, Paul D. Wilcox, and Anthony J. Croxford. A deep learning based methodology for artefact identification and suppression with application to ultrasonic images. *NDT & E International*, 126:102575, 2022.

[13] Johannes Vrana, Norbert Meyendorf, Nathan Ida, and Ripudaman (Ripi) Singh. *Introduction to NDE 4.0*, pages 1–28. Springer International Publishing, Cham, 2021.

[14] Marija Bertovic and Iikka Virkkunen. *NDE 4.0: New Paradigm for the NDE Inspection Personnel*, pages 239–269. Springer International Publishing, Cham, 2022.

[15] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2014.

[16] Charles Annis. Mil-hdbk-1823a, nondestructive evaluation system reliability assessment. Technical report, 2009.