A Case for Unsupervised Defect Detection in Manufacturing

Felix Viberg[1,2], Jonas Nordqvist[2], Morgan Ericsson[2], Anton Kaiser[3],
Martin Kroon[2], Welf Löwe[2], Marcus Nilsson[3], and Fredrik Sandberg[4]

***Abstract*—We compare a supervised and unsupervised approach to automated defect detection in high-performance manufacturing. We employ a U-Net convolutional neural network for the supervised approach and a normalizing flow model for the unsupervised approach. The experiments are conducted on two production datasets: bearing rings and coupling links. Our results demonstrate the initial effectiveness of both approaches in detecting defects, with the unsupervised methodology achieving competitive results compared to the state-of-the-art performance of a supervised segmentation model.**

I. INTRODUCTION

Automated defect detection is crucial for high-performance manufacturing as it reduces waste and labor costs while improving efficiency and ergonomics. Traditional supervised methods require large labeled datasets, which are challenging and time-consuming to obtain. This study investigates and compares supervised and unsupervised defect detection approaches in manufacturing components, aiming to offer a more efficient solution for defect detection.

A. Problem Statement

In high-performance manufacturing, automated defect detection faces challenges due to the low likelihood that good processes generate defective outcomes, leading to limited knowledge of error mode distributions. Furthermore, training state-of-the-art segmentation models to address this issue requires creating masks from large amounts of data for effective learning. This study aims to address these challenges and improve the efficiency of automated defect detection in the manufacturing domain using an unsupervised learning method.

B. Related Work

In computer vision tasks, supervised learning approaches often rely on convolutional neural networks (CNNs) such as U-Net [1] for segmentation tasks in manufacturing domains, including sheet steel surface inspection [2], semiconductor wafer defect detection [3], and weld defect detection [4]. These approaches typically require large amounts of labeled data.

Unsupervised learning approaches, such as clustering methods [5, 6], autoencoders [7], and variational autoencoders [8], aim to detect defects without labeled data. Recently, normalizing flows (see *e.g.* [9] and references therein) have emerged for defect detection in manufacturing data [10, 11].

*Supported by Vinnova (grant 2022-03018)
[1] SKF Sverige AB
[2] Linnæus University
[3] Gimic AB
[4] Gunnebo Industrier AB

II. METHOD

To address the limitations of supervised models, we explore the feasibility of using unsupervised methods for visual inspection of manufacturing components, enabling anomaly detection without the need for data labeling. Our approach builds on the experiments of [11], utilizing `EfficientNet-B5` [12] as a feature extractor and defining a multi-scale normalizing flows architecture.

A. Supervised and Unsupervised – Parallel approaches

For the supervised approach, we employ a U-Net architecture [1], a specialized CNN tailored for biomedical image segmentation tasks. The U-Net's design consists of an encoder path and a decoder path, with skip connections linking the corresponding layers from both paths. This structure enables the network to efficiently capture both local and global context, effectively integrating high-resolution features and contextual information to generate accurate segmentation results.

To detect defects using unsupervised learning, we employ the normalizing flow approach described in [11]. Normalizing flow is a technique used to learn a transformation of a potentially intricate (target) distribution into a (base) distribution with a known density (*e.g.*, standard Gaussian) using compositions of bijective transformations [9]. In our case, the target distributions are high-dimensional data distributions of images. The flow can be used for density evaluation in the target distribution. Outliers in the target distribution can be detected by computing likelihoods in the base distribution. The main assumption is that defects in the target distribution are outliers and will thus be detected as outliers in the base distribution.

B. Datasets

The experiment focuses on two production datasets collected from the production environment at two manufacturing sites.

1) Bearing Ring (BR): The dataset consists of high-resolution images of bearing ring faces, captured by two line-scan cameras as the ring rotates around its central axis, representing the faces as rectangular regions. To simulate authentic defects, damage was inflicted on the faces prior to heat treatment and grinding. 90% of the dataset (5511 patches) is used for training, and the remaining 10% is held out for validation and testing purposes.

2) Coupling links (CL): The dataset consists of medium-resolution images of blasted forged steel coupling links captured using an area-scan camera as the hook was translated and rotated by an industrial robot. Damaged pieces were collected using the regular manual inspection process in the production environment over an extended time frame. In the training set, there are 3424 images of non-defect hooks. The validation set is comprised of 568 images of which 165 contains a defect.

C. EXPERIMENT

1) Training Procedure: Supervised: The supervised U-Net model is trained using Binary Cross Entropy Loss in conjunction with Dice Loss [13]. The model is trained on pairs of 256-by-256 grayscale image and mask patches, derived from their full-resolution counterparts with a 50% overlap. Additionally, a set of transforms is used to augment the dataset, including random vertical and horizontal flipping, random rotation within 10 degrees, random scaling within 10%, and random translations within 10%. The weights are optimized with the Adam optimizer with a scheduled learning rate from 10^{-4} to 10^{-6} and decreasing for five non-improving epochs, β = (0.9, 0.999), and weight decay of 10^{-8}. A component is deemed defective if its predicted mask values exceed the predetermined 0.5 threshold.

2) Training Procedure: Unsupervised: We follow the approach in [11] for the unsupervised training procedure. The trained model is used to evaluate the likelihood of new samples in the target distribution. Low likelihood samples are considered outliers, potentially indicating defects. The likelihood threshold value is determined through a process of validation, analyzing the distribution of likelihood scores on the held-out subset of data and selecting an appropriate threshold that balances the trade-off between false positives and true positives.

III. RESULTS

Our experimental results are presented in Table I. The flow yields a negative log likelihood, and by setting a threshold, we can compute the model's precision and recall. The threshold is set to maximize the accuracy and is computed on the final model after training. The results are averaged over five training sequences. We further restrict our results to be measured at image level rather than pixel level for both model types.

In Figure 1, we illustrate the evaluation score trajectories of the unsupervised models across five training sequence runs. Each model is evaluated after 75 batches, and we register the current AUC-ROC. The bolded line is the average over all training sequences and individual training sequences are presented as traces. On top in each plot is the supervised AUC-ROC baseline based on our U-net model for the given dataset.

TABLE I: Precision and recall for supervised and unsupervised models on Bearing Rings (BR) and Coupling links (CL).

	Precision	Recall	AUC-ROC
Supervised (BR)	**0.939**	**0.914**	**0.928**
Unsupervised (BR)	0.829	0.804	0.861
Supervised (CL)	**0.981**	0.833	**0.912**
Unsupervised (CL)	0.600	**0.935**	0.859

IV. CONCLUSIONS AND FUTURE WORK

We compared supervised (U-Net) and unsupervised (normalizing flow) approaches for automated defect detection in manufacturing. While the supervised approach outperforms the unsupervised one in AUC-ROC, the latter offers the significant advantage of not requiring labeled data, which is valuable in settings where defects are rare and labeling costly. Although not mature, unsupervised learning methods show promising potential for defect detection in manufacturing, providing a more data-efficient alternative in real-world manufacturing environments. For further improvements, we plan to explore alternative feature extraction methods specifically tailored to industrial data. This approach could potentially enhance performance and generalization, as pretrained feature extractors are not optimized for the domain of interest. Currently, the inference time for the normalizing flow is much larger than for the supervised counterpart in our work. This is also a subject for further study.

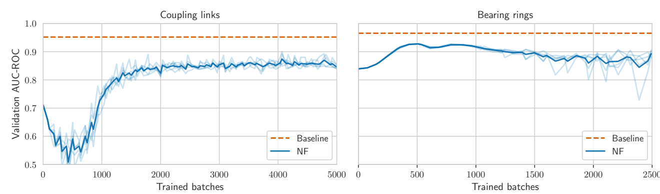

Fig. 1: Evaluation of five training sequences of the NF model.

REFERENCES

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," 2015, arXiv:1505.04597v1.

[2] R. Neven and T. Goedemé, "A multi-branch u-net for steel surface defect type and severity segmentation," *Metals (Basel)*, vol. 11, no. 6, p. 870, 2021.

[3] J. Cha and J. Jeong, "Improved U-Net with residual attention block for mixed-defect wafer maps," vol. 12, no. 4, p. 2209, 2022.

[4] L. Yang, H. Wang, B. Huo, F. Li, and Y. Liu, "An automatic welding defect location algorithm based on deep learning," *NDT & E international : independent nondestructive testing and evaluation*, vol. 120, p. 102435, 2021.

[5] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

[6] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *2nd ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, 1996, pp. 226–231.

[7] C.-Y. Liou, W.-C. Cheng, J.-W. Liou, and D.-R. Liou, "Autoencoder for words," *Neurocomputing*, vol. 139, pp. 84–96, 2014.

[8] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," 2022.

[9] G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed, and B. Lakshminarayanan, "Normalizing flows for probabilistic modeling and inference," *Journal of Machine Learning Research*, vol. 22, no. 57, pp. 1–64, 2021.

[10] M. Rudolph, B. Wandt, and B. Rosenhahn, "Same same but DifferNet: Semi-supervised defect detection with normalizing flows," in *Winter Conference on Applications of Computer Vision (WACV)*, Jan. 2021.

[11] M. Rudolph, T. Wehrbein, B. Rosenhahn, and B. Wandt, "Fully convolutional cross-scale-flows for image-based defect detection," in *Winter Conference on Applications of Computer Vision (WACV)*, Jan. 2022.

[12] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," 2020.

[13] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer International Publishing, 2017, pp. 240–248.