Deep Learning for Anomaly Detection in Industrial Manufacturing: A Study of Predictive Maintenance and Quality Control

Aaron Steiner

University of Mannheim, 68131 Mannheim, Germany
aaron.vitus.steiner@students.uni-mannheim.de

Abstract. This paper presents a comprehensive study on the application of anomaly detection in industrial settings, with a particular focus on the manufacturing sector. The paper discusses the role of predictive maintenance and quality control, particularly vision-based approaches, in enhancing operational efficiency and reducing economic losses due to unforeseen operational interruptions. It presents two instances where anomaly detection is actively being implemented in the industrial sector and explores various algorithms currently employed in this domain, including statistical methods and deep learning approaches. The study concludes with an examination of an experiment conducted by Zipfer et al. in an automotive manufacturing environment, highlighting the potential of deep learning methods in real-world industrial settings.

Keywords: Anomaly Detection · Deep Learning Methods · Industrial Applications

1 Introduction

The paradigm known as "Industry 4.0" represents a suite of technological advancements transforming the industrial landscape, with the goal of creating smart factories. This vision is realized through the integration of Internet of Things (IoT) devices, cyber-physical systems, and other interconnected cutting-edge technologies. These advanced tools, with their inherent intelligence and adaptability, have the potential to foster a more sophisticated and effective production process [16].

The enhancements in efficiency brought about by these technological innovations are multifaceted, with predictive maintenance and quality control being prime examples. Predictive maintenance is a process that assesses the likelihood of a machine component becoming dysfunctional in the near future, enabling preemptive replacement of the component and thus avoiding potential failure and associated downtime [8].

On the other hand, quality control, particularly vision-based approaches, leverages advanced imaging and machine learning techniques to detect anomalies and defects in products during the manufacturing process. This allows for real-time correction and improvement, ensuring the production of high-quality goods and reducing waste [17].

Current estimates suggest that unforeseen operational interruptions result in an annual financial deficit of approximately 50 billion dollars in the global industry [8]. Therefore, the advancements facilitated by Industry 4.0 promise not only enhanced efficiency in the production process but also a significant reduction in these economic losses.

Furthermore, the predictive capabilities offered by these technologies extend beyond maintenance and quality control, encompassing predictive analysis for supply chain management, demand forecasting, and product lifecycle management. This shift towards a more proactive approach could potentially revolutionize the industrial sector, ushering in an era of unprecedented productivity and efficiency [15].

The structure of this paper is as follows: Initially, we will present two instances where anomaly detection is actively being implemented in the industrial sector. Subsequently, we will delve into the exploration of algorithms currently employed in this domain. A brief overview of statistical methods will be provided, followed by an in-depth discussion on deep learning approaches, which are predominantly utilized for vision tasks in anomaly detection. These methods form the foundation for the experiment conducted by Zipfer et al. in an automotive manufacturing environment. This experiment is particularly noteworthy as it offers a unique perspective, given that most approaches are typically tested in laboratory settings, thereby overlooking the challenges encountered in real-world production environments. Finally, we will consolidate our observations, draw conclusions, and present our findings.

2 Anomaly Detection

2.1 Predictive maintenance

The evolution of maintenance paradigms within the manufacturing industry over the past century has been marked by significant transformations, moving from reactive to predictive models.

In the early stages, maintenance was primarily reactive in nature, implying that repairs were executed post the occurrence of machine failure. This mode of maintenance, however, bore notable drawbacks. Primarily, during a component failure, machines frequently suffered additional damages, leading to protracted and unforeseen periods of downtime. This unpredictability posed challenges to accurate production planning [3].

Subsequently, the paradigm shifted towards preventative maintenance, characterized by the establishment of service intervals for diverse components within a machine. While this approach enhanced the predictability of the maintenance schedule and thus improved planning capabilities, it necessitated frequent maintenance operations. An inherent characteristic of this approach is that it often

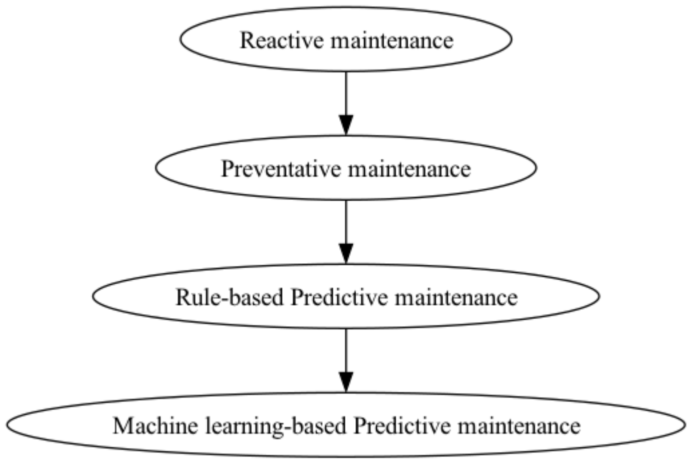

Fig. 1. Evolution of Maintenance in Manufacturing Industry

leads to the premature replacement of components that are still in optimal working condition and could have contributed to additional production cycles [3].

The onset of sensor technology heralded the advent of the initial version of predictive maintenance, known as rule-based maintenance. This technique utilized a set of pre-defined, hard-coded rules which, upon activation, triggered an alert instructing the operator to replace the respective component. Despite being an advancement over previous methods, this approach presented limitations due to its rigid and non-adaptive nature, which was not conducive to accommodating varying usage patterns [3].

This led to the emergence of the current iteration of predictive maintenance, which leverages machine learning-based methodologies to anticipate upcoming component failures. This proactive approach allows for maintenance operations to be performed prior to the actual failure, enhancing operational efficiency and minimizing unexpected downtimes. This paradigm shift in maintenance strategies, facilitated by the advent of sophisticated technologies, marks a significant milestone in the evolution of the manufacturing industry [3].

2.2 Quality control

Quality, as a parameter, holds a pivotal role across diverse sectors, yet it assumes an exceptional significance within the manufacturing industry due to its direct implications for waste reduction, reliability, and enhancement of production efficiency. The attainment of superior quality not only fulfills customer expectations, but also ensures compliance with legal standards, thus mitigating the potential ramifications associated with subpar quality outcomes. Quality Management (QM), therefore, assumes a crucial role in the manufacturing landscape, tasked with the continuous improvement and rigorous monitoring of quality benchmark [18].

Visual inspections are a prevalent non-destructive technique in product quality assurance. Given the direct relationship between visual integrity and product quality, these inspections play an important role, even though a visually perfect product could still possess hidden defects. In the past, inspections often relied on human operators manually checking product integrity, a process that was labor-intensive and potentially error-prone due to worker fatigue. This method, while detailed, presented challenges concerning consistency and efficiency [2].

In response to these challenges, modern visual inspection practices have incorporated algorithmic solutions from the field of anomaly detection for defect detection. Utilizing sophisticated image processing and machine learning algorithms, these automated systems significantly reduce the need for manual labor and enhance the consistency and accuracy of inspections. Moreover, these algorithmic solutions can operate continuously for extended periods without performance degradation, improving the reliability of the inspection process. They also enable real-time defect detection, which allows for prompt corrective actions. Therefore, the integration of algorithmic solutions within visual inspections has notably advanced the quality assurance mechanisms within the manufacturing industry [17].

One of the principal challenges in the domain of visual quality control pertains to the acquisition of a sufficient quantity of fault-related data, which is especially critical for supervised learning approaches, as will be explained in Section 3. The availability of such negative examples during the training phase is crucial for these models to effectively discern between intact and defective components [7].

The collection of faulty examples is particularly challenging due to the high proficiency of modern production processes that yield a minimal number of defective products, making them a rare occurrence. This scarcity, in turn, escalates the cost and complexity of the data collection process. While the intentional production of defective samples could be a potential solution, it is often associated with substantial time and financial investment. Further, the validity of artificially induced defects presents a dilemma, as it is uncertain whether such simulated defects accurately mirror the anomalies naturally manifested in the production process. Hence, assessing their relevance in anomaly detection becomes a complex task [17].

In response to these obstacles, unsupervised learning models have been proposed as an alternative solution. As detailed in Section 3, these models, which do not rely on labeled defective instances for training, are capable of identifying anomalies based on inherent data patterns. Thus, they offer a promising approach to navigate the limitations associated with the dearth of naturally occurring faulty examples in the data collection process [17].

3 Review of Existing Anomaly Detection Algorithms

Anomaly detection within the Internet of Things (IoT) can be categorized using various criteria, as suggested by researchers such as Fahim and Sillitti [5]. One common approach to classification is by the detection methods employed. Alternatively, one can categorize based on the applications or purposes of anomaly detection, such as constructive applications, malicious intent identification, or data cleaning. Another classification criterion is the type of anomaly - point anomalies, contextual anomalies, or collective anomalies. Lastly, the latency of the detection process can also serve as a distinguishing factor [5]. However, the focus of this section will be on presenting an overview of different anomaly detection methods, with a particular emphasis on statistical, and machine-learning approaches.

3.1 Statistical methods

In the realm of anomaly detection, statistical methods can be classified into two distinct groups based on their underlying assumptions. The first group consists of parametric algorithms, which assume a specific probability distribution for the data, often assuming a normal (Gaussian) distribution. These parametric models make certain assumptions about the data's statistical properties, such as mean and variance, and use these assumptions to identify anomalies. On the other hand, the second group comprises non-parametric algorithms that do not rely on explicit assumptions about the data's distribution. These non-parametric methods offer more flexibility and can be suitable for datasets that do not adhere to specific distribution assumptions. Two common examples in anomaly detection approaches are histogram-based methods and kernel-based functions [14]. According to Samara et. al. histogram and kernel, functions are the most often used methods in the field of anomaly detection [13].

Kernel functions: Kernel functions are a class of algorithms for pattern analysis, used for various applications such as outlier detection, clustering, and classification. They are used to transform the input data into a high-dimensional feature space where the data can be analyzed more effectively. The kernel function measures the similarity between two data points in this feature space. The choice of the kernel function depends on the nature of the data and the specific task at hand. For instance, in the context of outlier detection, a Gaussian kernel function might be used to model the probability density function of the data, enabling the identification of data points that deviate significantly from the norm [13].

Histograms: Histograms, on the other hand, are graphical representations that organize a group of data points into a specified range. In the context of outlier detection, histograms can be used to visualize the distribution of data, making it easier to identify values that lie far from the majority of the data. Histograms are particularly useful in univariate outlier detection, where the focus is on a single attribute or variable. The bins of the histogram represent the range of values, and the height of each bin corresponds to the frequency of data points within that range. Outliers are often represented by bins with significantly lower heights compared to the others [13].

In summary, both kernel functions and histograms can be used in the process of outlier detection. Kernel functions can be used to transform the data into a space where the outliers become more apparent, while histograms can provide a visual representation of the data distribution, making it easier to identify potential outliers [13].

3.2 Neural networks and deep learning architectures

Convolutional neural networks: Convolutional Neural Networks (CNNs) can be considered a specialized subset of general neural networks. These networks utilize a mathematical operation known as convolution, in contrast to the standard matrix multiplication typically used in fully connected neural networks. Convolution is a mathematical operation that involves combining two functions to produce a third. In the context of CNNs, convolution is utilized as an efficient mechanism for feature extraction. This operation is particularly effective for data that is structured in grid form. Images, for instance, are generally composed of a two-dimensional grid of pixels. However, convolution can also be applied to time-series data, which may be conceptualized as a one-dimensional grid with measurements taken at regular intervals [6].

The architecture of CNNs allows for the extraction of features at different levels of abstraction. In the context of image processing tasks, these features might include elementary components such as edges or corners in the initial layers. As one progresses deeper into the network, these primitive features are combined and recombined to form more complex, abstract features. This hierarchical feature learning allows the CNN to capture a surprisingly high level of detail in the input data. CNNs have demonstrated exceptional performance in a variety of image-related tasks. Their ability to automatically learn and extract a hierarchy of features from raw pixel data has made them an essential tool in the field of computer vision. The inherent structure of the convolution operation also confers a degree of translation invariance to the model, making CNNs robust to slight changes in the position or orientation of features within the input data. This combination of attributes has solidified the CNN's role as a cornerstone of modern machine learning, particularly in the realms of image recognition, segmentation, and classification [6].

Generative adversarial neural networks: GANs, introduced by Goodfellow et al. in 2014, have gained significant attention in the field of machine learning due to their unique ability to generate realistic data. The architecture of GANs is composed of two deep neural network models: a generative model (G) and a discriminator model (D). The generative model is tasked with learning the distribution of the training data and generating new samples from this learned distribution. The discriminator model, on the other hand, is designed to determine whether a given sample originates from the training data or was produced by the generative model [12].

The training process of GANs is characterized by an adversarial relationship between the generator and the discriminator. This adversarial process is akin to a zero-sum game, where the improvement of one model corresponds to an equal degradation in the performance of the other. The ultimate goal of this adversarial training is to reach a Nash equilibrium, a state where neither model can improve without the other model deteriorating. At this equilibrium, the discriminator model is unable to distinguish between real data and data generated by the generator model, resulting in a random guess about the origin of the input data [12].

In the context of anomaly detection, GANs present two significant advantages. Firstly, they can generate data that is hard to acquire, such as anomalous data points. Secondly, they can learn the distribution of normal operating data, thereby acting as an anomaly or outlier detector. This dual functionality of GANs has led to their application in two main ways: GAN-assisted anomaly detection and GAN-based anomaly detection [12].

In conclusion, GANs offer a promising approach to anomaly detection due to their ability to generate realistic data and learn data distributions. Their application spans various domains, and their use in both data augmentation and representation learning provides flexibility in addressing the challenges of anomaly detection. However, further research is needed to fully exploit the potential of GANs in this field [12].

Autoencoders: Autoencoders are neural networks that are trained to attempt to copy their input to their output. They consist of two main parts: an encoder function $\mathrm{h} = \mathrm{f}(\mathrm{x})$ and a decoder that produces a reconstruction $\mathrm{r} = \mathrm{g}(\mathrm{h})$. The encoder function compresses the input data and produces a code, while the decoder reconstructs the input data from this code. The aim of an autoencoder is to minimize the difference between the original input and its reconstruction. Autoencoders are designed to be unable to learn to copy perfectly. They are restricted in ways that allow them to copy only approximately, and to copy only input that resembles the training data. Because the model is forced to prioritize which aspects of the input should be copied, it often learns useful properties of the data [6].

Autoencoders can be regularized to learn useful features. One way to obtain useful features from the autoencoder is to constrain h to have a smaller dimension than x. An autoencoder whose code dimension is less than the input dimension is called undercomplete. Learning an undercomplete representation forces the autoencoder to capture the most salient features of the training data [6].

In the context of anomaly detection, autoencoders have been used effectively due to their ability to reconstruct normal data accurately while struggling to do so for anomalous data. The principle is simple: train the autoencoder on normal data so that it learns to reconstruct it accurately, then, when an anomalous data point is input, the autoencoder will likely reconstruct it poorly. This reconstruction error can then be used as an anomaly score, with higher scores indicating a higher likelihood of the data point being an anomaly [10].

In the paper "Deep Learning for Anomaly Detection: A Review" by Pang et al. the authors discuss the use of autoencoders for anomaly detection. They highlight that autoencoders are part of a broader category of methods that aim to learn feature representations of normality. These methods are designed to learn expressive representations of normal data, which can then be used to identify anomalies based on their deviation from these learned representations [10].

The authors also discuss the use of different types of autoencoders for anomaly detection, including variational autoencoders and denoising autoencoders. These variants introduce additional constraints to the encoding and decoding process, such as forcing the code to follow a specific distribution or adding noise to the input data to improve the robustness of the learned representations [10].

In summary, autoencoders provide a powerful tool for anomaly detection by learning to reconstruct normal data and using the reconstruction error as an anomaly score. Their ability to learn complex data representations and their flexibility to incorporate additional constraints make them a popular choice for anomaly detection tasks [10].

4 Anomaly detection in automotive manufacturing

4.1 Study introduction

The experimental framework delineated by Zipfer et al. is predicated on a collaborative endeavor with a prominent European automobile manufacturer. A subsidiary of the manufacturer is in the process of developing a cloud-based infrastructure tailored for industrial applications, with a particular emphasis on the automation of quality assurance processes. One such application pertains to the inspection of registration-relevant labels affixed to vehicles during the production phase. These labels are scrutinized for their presence, layout, position, content, and reading direction, with the overarching objective being the development of models capable of detecting anomalies in these labels [17].

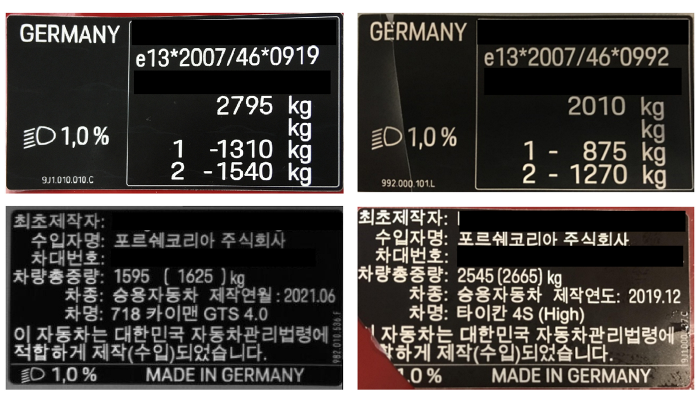

Fig. 2. Example of a VIN label [17]

The labels are photographed in two distinct scenarios: on the assembly line via stationary cameras, and during rework using mobile devices. The images thus captured present a formidable challenge owing to the variability in lighting conditions and the extensive diversity of the training data, as the Vehicle Identification Number (VIN) labels exhibit a multitude of country-specific layouts. Furthermore, the absence of an international standard for these labels necessitates that the models either possess robust generalization capabilities or that separate models be trained for each individual country. The criticality of the accuracy and integrity of these VIN labels cannot be overstated, as any discrepancies or damage could potentially impede the import or registration of the vehicle in another country, thereby posing a significant financial risk to the manufacturer [17].

In the course of the experiments, a diverse collection of real-world images encapsulating VIN labels from various vehicles was amassed. This collection comprised images of both damaged and undamaged labels, captured using stationary cameras and smartphone cameras. However, given the limited occurrence of damage to VIN labels in the production process, anomaly cases had to be artificially constructed. This was achieved by intentionally damaging VIN labels, affixing them to test vehicles, and subsequently photographing them. It remains unclear whether these photographs were taken at the production line or at a different location. This process culminated in the compilation of a dataset consisting of 2703 images of undamaged VIN labels and 970 images of damaged VIN labels [17].

4.2 Evaluated models

In the study conducted by Zipfer et al., four distinct models were selected for evaluation: Skip-GANomaly, PaDiM, PatchCore, and Auto-Classifier. The selection of these models was predicated on an extensive literature review. It is noteworthy that three out of the four models represent unsupervised approaches, while the Auto-Classifier serves as a supervised performance benchmark. Although the unsupervised models do not necessitate the inclusion of examples of damaged VINs during the training process, these examples are nonetheless integral to the subsequent evaluation process [17].

1. Skip-GANomaly, an unsupervised model, utilizes a reconstruction-based methodology. The training phase is conducted solely on normal (undamaged) images, while the testing phase incorporates both normal and anomalous (damaged) images. The computation of the anomaly score is achieved through the weighted sum of the reconstruction loss and the latent-representation loss [1].
2. PaDiM is an unsupervised model that implements a representation-based strategy. It does not depend on the reconstruction loss of an image. Instead, it leverages the similarity of embedding vectors. An embedding vector is characterized as the concatenated feature maps extracted from various hidden layers of a Convolutional Neural Network (CNN) [4].
3. PatchCore represents another unsupervised model that adheres to a representation-based approach. It exploits a pre-trained CNN backbone network to derive embedding vectors. Initially, these vectors are stored in a structure known as

a memory bank. To reduce computational load, the memory bank undergoes subsampling through the application of a k-center-greedy algorithm [11].

4. Finally, Auto-Classifier is a supervised model that integrates a CNN-fusion approach with an Automated Machine Learning (AutoML) procedure. Several pre-trained CNN architectures are subjected to retraining. The test results from all CNN architectures are subsequently amalgamated, producing a single classification result that can be interpreted as an anomaly score [9].

4.3 Achieved results

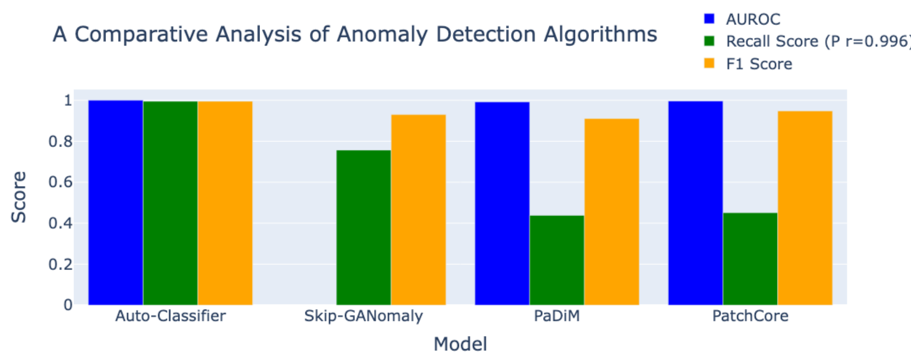

Fig. 3. A Comparative Analysis of Anomaly Detection Algorithms[1]

In the comparative analysis of the models, the Auto-Classifier emerged as the superior model in distinguishing between damaged and intact Vehicle Identification Number (VIN) labels. Remarkably, depending on the seed, it achieved an F1 score of up to 1, indicating flawless performance on the test set. The score dropped by a maximum of 0.01 with a different seed. Among the deep learning models, PatchCore demonstrated the best performance, achieving an F1 score of 0.95. The performance of PaDiM varied depending on the backbone used, but it still scored in the high 80s to low 90s range. The least effective model was Skip-GANomaly, which achieved an F1 score of just 0.76. In terms of inference time, i.e., the time the model requires to evaluate a given image, PatchCore was notably slower, taking almost 2 seconds to make a decision about an image [17].

Subsequent experiments were conducted to enhance the scores of the deep learning models. The supervised approach was excluded from these experiments as it had already achieved a perfect score. The first experiment aimed to reduce the variance in the images by aligning every image in the reading direction. This adjustment predictably improved the scores of all models, highlighting the

[1] Data based on [17]

importance of such a preparatory step that could otherwise be easily overlooked [17].

In the second experiment, a distinction was made between different images, with stationary images scored independently from mobile ones. As expected, the stable images were easier to evaluate due to their lower variance. This was reflected in all models achieving lower scores on the mobile images compared to the stationary ones. However, the impact was not uniform across the different models. In fact, Skip-GANomaly's F1 score dropped almost 0.4 when compared to the stationary images, while the other models proved to be more robust. PaDiM only lost 0.02 in F1 with PatchCore dropping 0.03, with both models maintaining an impressive F1 score of about 0.95 on the mobile images. Demonstrating in their findings that representation-based models seemed to outperform the reconstruction-based approach (Skip-GANomaly) [17].

4.4 Critic

The paper by Zipfer et al., which investigates the application of various machine learning models for anomaly detection in vehicle identification number (VIN) labels, provides valuable insights, particularly due to its real-world production setting. However, there are several noteworthy aspects that warrant further scrutiny.

The Auto-Classifier model's achievement of a perfect F1 score of 1.0 is unusual in real-world applications. While not impossible, such a score raises questions about the complexity and diversity of the test set. A perfect score could potentially indicate that the test set may not be sufficiently diverse or challenging to fully evaluate the model's performance. It could also suggest overfitting of the model to the training data, which would limit its ability to generalize to new, unseen data. One possible explanation for this score could be the artificial creation of damaged VIN examples due to a lack of naturally occurring instances. While the authors seem to have invested significant effort in this process, it is impossible to determine whether these artificially created examples share some common traits that the model could have exploited. This is due to the data not being openly available due to restrictions imposed by the car manufacturer.

Although the authors do not directly address this issue, they suggest in their discussion that unsupervised models might perform better in a production setting as the examples of damaged VINs could change over time due to alterations in the manufacturing process. This would not pose a problem for the deep learning models [17].

Finally, the lack of a benchmark makes it difficult to evaluate the implications of their results. As previously explained, the supervised benchmark has some issues. While the authors provide some insights into the manufacturer's requirements, namely a precision score of above 0.996 [17], it would have been insightful to understand the performance of the solution currently being applied for this task. If the company currently does not have a model in place for this task, which seems likely due to some comments in the paper (namely, that they did not have examples of broken VINs), this could have been made clearer.

5 Conclusion

In the context of this paper, we have examined two instances where anomaly detection is actively being implemented in the industrial sector. We have also explored various algorithms currently employed in this domain, with a focus on deep learning approaches, which are predominantly utilized for vision tasks in anomaly detection. The experiment conducted by Zipfer et al. in an automotive manufacturing environment is particularly noteworthy as it offers a unique perspective, given that most approaches are typically tested in laboratory settings, thereby overlooking the challenges encountered in real-world production environments.

The potential of deep learning methods in anomaly detection, particularly in the context of manufacturing, is significant. These methods, which do not require negative examples during training, offer a promising approach for a wide range of applications. The ability to learn from normal data and identify deviations from this norm makes these models highly adaptable and versatile. The results of the study by Zipfer et al. demonstrate the potential of deep learning methods in real-world industrial settings. The ability of these models to accurately detect anomalies in VIN labels, despite the variability in lighting conditions and label layouts, is a testament to their robustness and adaptability.

However, the application of these models is not without challenges. In certain domains, such as medicine, the situation is often reversed. There is an abundance of images depicting various medical conditions, but relatively few of healthy or normal states. This presents a unique challenge for anomaly detection models that are trained on normal data. While this does not preclude the use of these models in such domains, it does necessitate a different approach to training and model selection.

Looking forward, it is clear that deep learning methods will continue to play a pivotal role in anomaly detection. As these models continue to evolve and improve, they will undoubtedly find application in a wide range of domains, from manufacturing to medicine and beyond. However, it is crucial that we continue to critically evaluate these models and their performance in different settings, to ensure that they are being used effectively.

In summary, while deep learning methods offer a promising approach to anomaly detection, their successful application requires careful consideration of the specific characteristics and requirements of the domain in question. With the right approach to training and model selection, these methods have the potential to revolutionize anomaly detection across a wide range of industries.

References

1. Samet Akçay, Amir Atapour Abarghouei, and Toby P. Breckon. Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection. *CoRR*, abs/1901.08954, 2019.

2. Danilo Avola, Marco Cascio, Luigi Cinque, Alessio Fagioli, Gian Luca Foresti, Marco Raoul Marini, and Fabrizio Rossi. Real-time deep learning method for automated detection and localization of structural defects in manufactured products. *Computers & Industrial Engineering*, 172:108512, 2022.
3. P Coandă, M Avram, and V Constantin. A state of the art of predictive maintenance techniques. *IOP Conference Series: Materials Science and Engineering*, 997(1):012039, dec 2020.
4. Thomas Defard, Aleksandr Setkov, Angelique Loesch, and Romaric Audigier. Padim: a patch distribution modeling framework for anomaly detection and localization. *CoRR*, abs/2011.08785, 2020.
5. Muhammad Fahim and Alberto Sillitti. Anomaly detection, analysis and prediction techniques in iot environment: A systematic literature review. *IEEE Access*, 7:81664–81681, 2019.
6. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.
7. Christian Janiesch, Patrick Zschech, and Kai Heinrich. Machine learning and deep learning. *Electronic Markets*, 31(3):685–695, 2021.
8. Pooja Kamat and Rekha Sugandhi. Anomaly detection for predictive maintenance in industry 4.0- a survey. *E3S Web of Conferences*, 170:02007, 2020.
9. Vasco Lopes and Luís A Alexandre. Auto-classifier: A robust defect detector based on an automl head. In *Neural Information Processing: 27th International Conference, ICONIP 2020, Bangkok, Thailand, November 23–27, 2020, Proceedings, Part I*, pages 137–149. Springer, 2020.
10. Guansong Pang, Chunhua Shen, Longbing Cao, and Anton Van Den Hengel. Deep learning for anomaly detection. *ACM Computing Surveys*, 54(2):1–38, mar 2021.
11. Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, and Peter Gehler. Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14318–14328, 2022.
12. Mikael Sabuhi, Ming Zhou, Cor-Paul Bezemer, and Petr Musilek. Applications of generative adversarial networks in anomaly detection: A systematic literature review, 2021.
13. Mustafa Al Samara, Ismail Bennis, Abdelhafid Abouaissa, and Pascal Lorenz. A survey of outlier detection techniques in iot: Review and classification. *Journal of Sensor and Actuator Networks*, 11(1), 2022.
14. Durgesh Samariya and Amit Thakkar. A comprehensive survey of anomaly detection algorithms. *Annals of Data Science*, 10(3):829–850, 2023.
15. Mahya Seyedan and Fereshteh Mafakheri. Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, 7(1):53, 2020.
16. Ray Y. Zhong, Xun Xu, Eberhard Klotz, and Stephen T. Newman. Intelligent manufacturing in the context of industry 4.0: A review. *Engineering*, 3(5):616–630, 2017.
17. Justus Zipfel, Felix Verworner, Marco Fischer, Uwe Wieland, Mathias Kraus, and Patrick Zschech. Anomaly detection for industrial quality assurance: A comparative evaluation of unsupervised deep learning models. *Computers & Industrial Engineering*, 177:109045, 2023.
18. Avigdor Zonnenshain and Ron S. Kenett. Quality 4.0—the challenging future of quality engineering. *Quality Engineering*, 32(4):614–626, 2020.