POLITECNICO DI TORINO
Repository ISTITUZIONALE

Neuro-symbolic Empowered Denoising Diffusion Probabilistic Models for Real-time Anomaly Detection in Industry 4.0

Original
Neuro-symbolic Empowered Denoising Diffusion Probabilistic Models for Real-time Anomaly Detection in Industry 4.0 / Capogrosso, L., Mascolini, A., Girella, F., Skenderi, G., Gaiardelli, S., Dall'Ora, N., Ponzio, F., Fraccaroli, E., DI CATALDO, S., Vinco, S., Macii, E., Fummi, F., Cristani, M.. - (2023), pp. 1-4. (Forum on specification & Design Languages Turin, Italy 13-15 September 2023) [10.1109/FDL59689.2023.10272095].

Availability:
This version is available at: 11583/2982396 since: 2023-10-11T15:59:36Z

Publisher:
IEEE

Published
DOI:10.1109/FDL59689.2023.10272095

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

27 November 2026

Neuro-symbolic Empowered Denoising Diffusion Probabilistic Models for Real-time Anomaly Detection in Industry 4.0

Wild-and-Crazy-Idea Paper

Luigi Capogrosso*, Alessio Mascolini†, Federico Girella*, Geri Skenderi*, Sebastiano Gaiardelli*, Nicola Dall'Ora*, Francesco Ponzio‡, Enrico Fraccaroli§, Santa Di Cataldo†, Sara Vinco†, Enrico Macii‡, Franco Fummi*, Marco Cristani*

*Dept. of Engineering for Innovation Medicine, University of Verona, Italy, name.surname@univr.it

†Dept. of Control and Computer Engineering, Polytechnic of Turin, Italy, name.surname@polito.it

‡Dept. of Regional and Urban Studies and Planning, Polytechnic of Turin, Italy, name.surname@polito.it

§Dept. of Computer Science, University of North Carolina at Chapel Hill, USA, enrifrac@cs.unc.edu

***Abstract*—Industry 4.0 involves the integration of digital technologies, such as IoT, Big Data, and AI, into manufacturing and industrial processes to increase efficiency and productivity. As these technologies become more interconnected and interdependent, Industry 4.0 systems become more complex, which brings the difficulty of identifying and stopping anomalies that may cause disturbances in the manufacturing process. This paper aims to propose a diffusion-based model for real-time anomaly prediction in Industry 4.0 processes. Using a neuro-symbolic approach, we integrate industrial ontologies in the model, thereby adding formal knowledge on smart manufacturing. Finally, we propose a simple yet effective way of distilling diffusion models through Random Fourier Features for deployment on an embedded system for direct integration into the manufacturing process. To the best of our knowledge, this approach has never been explored before.**

***Index Terms*—Industry 4.0, Anomaly Detection, Diffusion Models, Neuro-symbolic AI, Knowledge Distillation**

I. Context and Motivation

The dawn of Industry 4.0 has has ushered a digital revolution in industrial processes, significantly increasing productivity, efficiency, and quality. This transformation is facilitated by the Industrial Internet of Things (IIoT) [1], which involves the interconnection of industrial devices, equipment, and systems through the Internet. The generation and collection of massive amounts of diverse data enabled by IIoT provide insights into various aspects of production, such as process optimization, quality control, and resource allocation. But, on the other hand, bring new challenges in data analysis and management.

This study was carried out within the PNRR research activities of the consortium iNEST (Interconnected North-Est Innovation Ecosystem) funded by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.5 – D.D. 1058 23/06/2022, ECS_00000043). This manuscript reflects only the Authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

979-8-3503-0737-5/23/$31.00 ©2023 IEEE

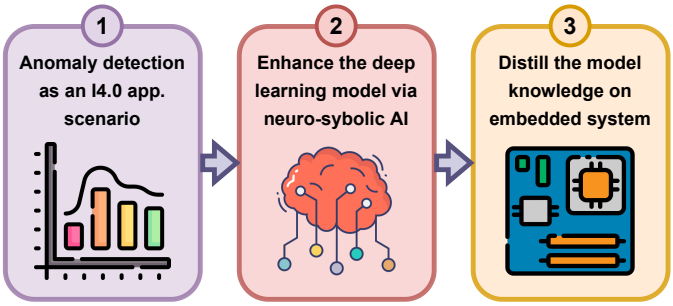

Fig. 1. The key topics of our proposal. *1)* Starting from the open problem of anomaly detection in an Industry 4.0 scenario, *2)* we propose to use a recent category of deep learning models (*i.e.*, diffusion models) to address this problem, by enhancing them with neuro-symbolic learning, which is one of the trending topics of the moment in the field of AI. *3)* Finally, we propose a distillation strategy, to transfer the knowledge into an embedded system for real-time usage on the product line.

One of the most critical tasks in this context is *anomaly detection*, which refers to the identification of irregular events or behaviors from the data. Malfunctioning equipment, human errors, external disturbances, or unforeseen circumstances can disrupt production, pose safety hazards, or lower product quality [2]. As such, the development of techniques able to accurately and timely detect anomalies is crucial for maintaining smooth and efficient operations, as well as ensuring the safety and well-being of workers.

Although remarkable improvements have been made in recent literature, especially in the field of deep learning , existing anomaly detection approaches still struggle to effectively model the behavior of complex machinery while maintaining their computational effort within feasible limits for real-time execution at the edge [3]. This motivation is the core of our research proposal.

II. Our Proposal

We want to enhance the productivity and safety of the production line by implementing a system that can effectively

detect irregular behaviors of the system, even when human attention is lacking or unavailable. Traditional methods for achieving this goal rely on purely symbolic approaches, as companies need assurance that the system will provide explainable predictions and will not demand extensive computing resources. However, due to the simplicity of these methods, they are often limited in their ability to recognize anomalies that rely on context or relationships between sensors. Another option is to rely on deep learning models, which can learn complex relationships in the sensed data. Unfortunately, they are often avoided due to their "black-box" nature and the amount of effort required to run even the smallest models in real-time. Furthermore, they might require more expensive hardware and higher power usage, without the guarantee to match or exceed the performance of symbolic approaches.

To address these concerns, this paper presents an innovative approach to the integration of a formally constrained diffusion model for anomaly detection with embedded systems for industrial applications, merging neuro-symbolic methods, formal constraints to enhance reliability and safety in industrial systems, and knowledge distillation. This innovative solution strives to improve monitoring capabilities without sacrificing the reliability and efficiency required in an Industry 4.0 production line setting. Figure 1 shows our proposed anomaly detection flow. What follows is a detailed explanation of the flow which, we hope, might pose the basis for further investigation and research in this area.

We approach anomaly detection as a task of Out-Of-Distribution (OOD) classification. Anomaly detection is the task of identifying examples in a dataset that are different or unusual compared to most other examples. Similarly, OOD classification is the task of identifying examples that do not belong to a known data distribution. Both tasks involve identifying examples dissimilar to the majority of the other examples, and both require the ability to identify patterns in the data that are significantly different from the norm. Inspired by the work of [4], we propose a Denoising Diffusion Probabilistic Model (DDPM) [5] as the backbone for our anomaly detection process. We take advantage of DDPMs' ability to understand a given distribution latent structure to detect OOD samples in our data via differences in signal reconstruction. The final goal of the DDPM will be to label the training data in a fully unsupervised way.

Our second contribution is exploiting ontologies to add formal and additional knowledge as properties that can be integrated within the deep learning model, to constrain the diffusion model to learn a data distribution that always respects the given logical axioms. We plan to do so by extending the methodology presented in [6], where authors present a formal model able to represent abstract capabilities, structure, and executable skills of a manufacturing system through so-called ontology design patterns based on industrial standards, *e.g.*, DIN 8580 [7], VDI 3682 [8], and VDI 2860 [9]. By enriching our network with domain-specific knowledge, we ensure the model starts from a formal understanding of the data and can detect the anomalies that a purely symbolic system should be able to detect. Combined with the Deep Generative Model (DGM), this allows the network to learn more complex rules and relationships, and thus more sophisticated anomalies.

Specifically, we plan to use a neural-symbolic AI [10] approach to support the learning phase of neural networks using the satisfaction of a first-order logic knowledge base as an objective, similarly to [11]. In our case, this is provided by the ontologies that formalize the smart manufacturing knowledge in an interoperable way. From the formal conceptualization given by the ontologies, we obtain the first-order logic knowledge base containing a set of axioms. At this point, we have some predicates, or functions appearing in these axioms that we want to learn, and some data available that we can use to learn the parameters of those symbols. The idea is to embed the logical axioms into the loss function of our diffusion model. The goal of our model then becomes finding solutions in the hypothesis space that maximally satisfy all the axioms contained in our knowledge base.

This hybrid method, which combines both data-driven and knowledge-based techniques, allows the neural network to discern the remaining nuances and intricacies of the system, which should ultimately enhance its overall performance and reliability. The described system is explainable by design, as it allows the user to know which sensors are reporting anomalous data, and which values the system would have considered acceptable.

Finally, we address the challenge of running the algorithm in real-time on the product line, which is in direct contrast with the computationally intensive nature of deep learning models, particularly diffusion models. In order to solve this, our third proposal consists of the use of Random Fourier Features (RFF) [12] to train a classifier on the binary labels obtained by our DDPM, effectively distilling its knowledge, without using a teacher-student learning paradigm [13], into a more lightweight and practically useful detector. RFF-based classifiers represent an extension of linear classifiers that project the data into a higher dimensional space, where the classes are more easily separable, in order to model arbitrary non-linear functions. By training a RFF model on the pseudo-labels created by the NeSy-DDPM, it is plausible to think that RFFs can be used to provide an optimal decision boundary in kernel space [14] for our proposed anomaly detection classifier. To sum up, the RFF must not capture the posterior inferred by the diffusion model per se, but they must be able to define features that support the (binary) label distribution generated by the diffusion-based-OOD [4]. Given a kernel function $K(x, y)$, we can represent it as an inner product in a higher-dimensional space via the feature map:

$$\phi(x) = \begin{bmatrix} \cos(w_1^T x + b_1) \\ \sin(w_1^T x + b_1) \\ \vdots \\ \cos(w_D^T x + b_D) \\ \sin(w_D^T x + b_D) \end{bmatrix}, \tag{1}$$

where D is the desired dimensionality of the feature space, and w_i and b_i are randomly generated parameters.

To train a linear classifier using RFF, we first compute the feature representation for each input data point using the above formula and then use it to predict the label previously assigned to that sample by our DDPM approach. Once the linear classifier is trained on these features, we can make predictions on new data points by computing their feature representation and then multiplying them by the learned weight vector w and adding a bias term b:

$$p(y_{pred}) = \frac{1}{1 + e^{-(w^T \phi(x_{new}) + b)}} \ . \tag{2}$$

From 2, the entire inference process requires only two matrix multiplications, making them extremely lightweight and quickly executable on low-end hardware.

In conclusion, our idea, shown in Figure 2, consists of a series of proposals in order to leverage the power of diffusion models for anomaly detection from time series data. The framework can capture the complex spatiotemporal dependencies between the sensors, handle large amounts of data in a rigorous way, and enable proactive anomaly detection and mitigation, thereby enhancing the reliability, safety, and efficiency of industrial processes. For all of these reasons, we think that our proposal can be a valuable tool for Industry 4.0.

III. Related Work

This section introduces the techniques we rely upon.

A. Diffusion models

Given observed samples x from a distribution of interest, the goal of a generative model is to learn to estimate its true data distribution $p(x)$. Once learned, we are able to use the learned model to evaluate the likelihood of newly observed data. DDPMs are parameterized Markov chains models, able to learn the latent structure of data by modeling the way data points diffuse through the latent space. They emerged in 2015 [15] as a powerful new family of deep generative models with a record-breaking performance in a broad range of applications, spanning from image synthesis and multimodal modeling to temporal analysis and natural language problem [16]. Referring to our problem, there are three principal works related to OOD with diffusion models.

The authors of [4] utilize DDPMs as denoising autoencoders, where the amount of noise applied externally controls the strength of the conditioning. They suggest employing DDPMs to reconstruct a noised input across various noise levels and leverage the resulting multidimensional reconstruction error to classify OOD inputs. The experiments demonstrate that the proposed DDPM-based technique surpasses both reconstruction-based methods and state-of-the-art generative approaches.

As previously stated, DGMs appear to be a natural choice for detecting OOD inputs. However, these models have been observed to assign higher probabilities or densities to OOD images than images from the training distribution. The work in [17] addresses this behavior and attributes it to model misestimation, suggesting it to be a more probable cause, rather than the misalignment between likelihood-based OOD detection and our distributions of interest. They also demonstrate how even slight estimation errors can result in OOD detection failures, which carries implications for future research in deep generative modeling and OOD detection. This work, although primarily focused on image inputs, has shown that DDPMs can be utilized to address this problem. This provides theoretical foundations for our idea, even though it may seem unconventional.

B. Neuro-symbolic AI

Neuro-symbolic AI refers to the combination of artificial neural networks with symbolic knowledge representation and reasoning techniques used in symbolic AI. This approach aims to overcome the limitations of traditional rule-based symbolic AI by incorporating both logical reasoning and statistical inference into the same model [10].

One of these studies is that of Zheng *et al*. [18], tackling the challenge of developing a conversational embodied agent capable of executing real-life tasks. Traditional symbolic methods suffer from scaling and generalization issues, while end-to-end deep learning models face data scarcity and high task complexity, and are often difficult to interpret. To take advantage of both approaches, the authors propose a neuro-symbolic commonsense reasoning framework for generalizable and interpretable conversational embodied agents.

Another recent work is the one by Huang *et al*. [19], in which they propose a neuro-symbolic approach that learns semantic representations by leveraging logic specifications that can capture rich spatial and temporal properties in video data.

Furthermore, Siyaev *et al*. [20] proposed a neuro-symbolic reasoning approach for interacting with 3D digital twins through natural language. This method can comprehend user requests and contexts to manipulate 3D components of digital twins and execute installations and removal procedures independently by reading maintenance manuals.

Overall, all these works demonstrate the potential of neuro-symbolic AI in various domains. The proposed approaches not only improve the accuracy of predictions but also provide valuable insights into the decision-making process, making them more transparent and trustworthy.

IV. Discussion

This paper outlines a novel idea that combines neuro-symbolic diffusion models and RFF for real-time anomaly detection. This section discusses the reasons why this idea is considered wild and crazy and the potential implications of successfully implementing such a model.

First and foremost, while the idea of integrating symbolic and neural systems has been explored in other contexts, such as in hybrid neuro-symbolic models, the concept of neuro-symbolic diffusion models is entirely new and has never been attempted before. This approach proposes to combine the strengths of symbolic and neural systems in a single model, potentially enabling a more robust representation of

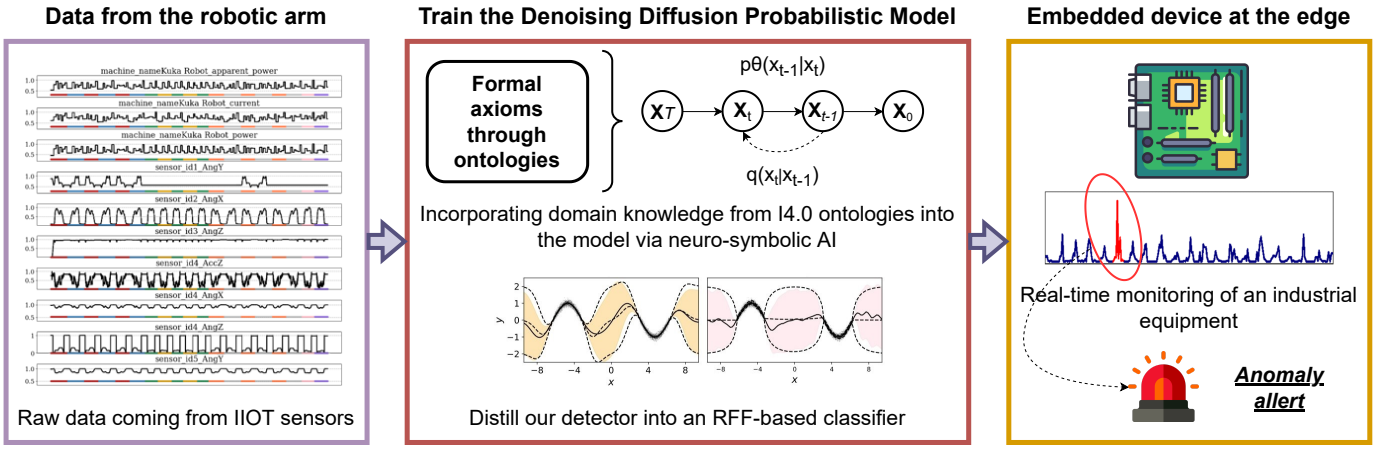

Fig. 2. Detailed overview of our proposal. The time series on which we want to train the diffusion model are acquired by the sensors of the robotic arm. Diffusion models represent the latest advancement in deep learning, and thus, are potentially capable of better performance compared to previous models. This model is enhanced through a neural-symbolic approach using the satisfaction of a first-order logic knowledge base as an objective, which, in our case, is provided by the ontologies that formalize the smart manufacturing knowledge in an interoperable way. Finally, we face the problem of running the algorithm in real-time on the product line effectively distilling its knowledge into an RFF-based classifier. As a result, the entire inference process requires only two matrix multiplications, which can be executed quickly on resource-constrained devices.

complex information. At the same time, it relies on domain-specific logical constraints to guide the DDPM during training, meaning that careful considerations on the formalisms used are needed by domain experts.

Secondly, while RFF classifiers have been used in other contexts, utilizing them as a means to distill DDPMs is unexplored and presents numerous challenges. That being said, their low computation complexity offers a promising avenue for deployment on embedded systems, allowing us to utilize the power of deep learning models even in resource-constrained environments.

The complexity of Industry 4.0 systems and the datasets they produce can present significant challenges for any anomaly detection approach. Anomaly detection is a critical task that demands high accuracy and reliability; as such, any approach utilized for this purpose must undergo thorough testing and validation to ensure its effectiveness. It is vital to consider whether the proposed approach can handle the complexity of these systems and datasets and whether it can scale to meet the demands of real-world applications.

Although it may seem unconventional at first glance, our approach is grounded in sound scientific principles and has the potential to significantly contribute to the field of smart manufacturing, offering a promising solution to the challenges posed by real-time anomaly detection in Industry 4.0.

References

[1] E. Sisinni, A. Saifullah, S. Han *et al.*, "Industrial internet of things: Challenges, opportunities, and directions," *IEEE transactions on industrial informatics*, vol. 14, no. 11, pp. 4724–4734, 2018.

[2] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[3] G. Pang, C. Shen, L. Cao *et al.*, "Deep learning for anomaly detection: A review," *ACM computing surveys (CSUR)*, vol. 54, no. 2, pp. 1–38, 2021.

[4] M. S. Graham, W. H. Pinaya, P.-D. Tudosiu *et al.*, "Denoising diffusion models for out-of-distribution detection," *arXiv preprint arXiv:2211.07740*, 2022.

[5] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6840–6851, 2020.

[6] A. Köcher, C. Hildebrandt, L. M. V. da Silva *et al.*, "A formal capability and skill model for use in plug and produce scenarios," in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, vol. 1. IEEE, 2020, pp. 1663–1670.

[7] "Industrialstandard-odp-din8580," https://github.com/hsu-aut/IndustrialStandard-ODP-DIN8580, 2023.

[8] "Industrialstandard-odp-vdi3682," https://github.com/hsu-aut/IndustrialStandard-ODP-VDI3682, 2023.

[9] "Industrialstandard-odp-vdi2860," https://github.com/hsu-aut/IndustrialStandard-ODP-VDI2860, 2023.

[10] M. K. Sarker, L. Zhou, A. Eberhart *et al.*, "Neuro-symbolic artificial intelligence," *AI Communications*, vol. 34, no. 3, pp. 197–209, 2021.

[11] S. Roychowdhury, M. Diligenti, and M. Gori, "Regularizing deep networks with prior knowledge: A constraint-based approach," *Knowledge-Based Systems*, vol. 222, p. 106989, 2021.

[12] A. Rahimi and B. Recht, "Random features for large-scale kernel machines," *Advances in neural information processing systems*, vol. 20, 2007.

[13] J. Gou, B. Yu, S. J. Maybank *et al.*, "Knowledge distillation: A survey," *International Journal of Computer Vision*, vol. 129, pp. 1789–1819, 2021.

[14] M.-N. Nguyen and N. A. Vien, "Scalable and interpretable one-class svms with deep learning and random fourier features," in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018, Dublin, Ireland, September 10–14, 2018, Proceedings, Part I 18*. Springer, 2019, pp. 157–172.

[15] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan *et al.*, "Deep unsupervised learning using nonequilibrium thermodynamics," in *International Conference on Machine Learning*. PMLR, 2015, pp. 2256–2265.

[16] L. Yang, Z. Zhang, Y. Song *et al.*, "Diffusion models: A comprehensive survey of methods and applications," *arXiv preprint arXiv:2209.00796*, 2022.

[17] L. Zhang, M. Goldstein, and R. Ranganath, "Understanding failures in out-of-distribution detection with deep generative models," in *International Conference on Machine Learning*. PMLR, 2021, pp. 12 427–12 436.

[18] K. Zheng, K. Zhou, J. Gu *et al.*, "Jarvis: A neuro-symbolic commonsense reasoning framework for conversational embodied agents," *arXiv preprint arXiv:2208.13266*, 2022.

[19] J. Huang, Z. Li, D. Jacobs *et al.*, "Laser: Neuro-symbolic learning of semantic video representations," *arXiv preprint arXiv:2304.07647*, 2023.

[20] A. Siyaev, D. Valiev, and G.-S. Jo, "Interaction with industrial digital twin using neuro-symbolic reasoning," *Sensors*, vol. 23, no. 3, p. 1729, 2023.