Human-Centric Proactive Quality Control in Industry5.0: The critical role of explainable AI

PaoloCatti
Laboratory for Manufacturing Systems & Automation
Department of Mechanical Engineering and Aeronautics, University of Patras
Rio, Patras, Greece
catti@lms.mech.upatras.gr

Emmanouil Bakopoulos
Laboratory for Manufacturing Systems & Automation
Department of Mechanical Engineering and Aeronautics, University of Patras
Rio, Patras, Greece
bakopoulos@lms.mech.upatras.gr

Anja Stipankov
F6S EU Tech Innovation Network
77 Camden Street Lowe, Saint Kevin's
Dublin, Ireland
anja@f6s.com

Natalia Cardona
F6S EU Tech Innovation Network
77 Camden Street Lowe, Saint Kevin's
Dublin, Ireland
natalia@f6s.com

Nikolaos Nikolakis
Laboratory for Manufacturing Systems & Automation
Department of Mechanical Engineering and Aeronautics, University of Patras
Rio, Patras, Greece
nikolakis@lms.mech.upatras.gr

Kosmas Alexopoulos
Laboratory for Manufacturing Systems & Automation
Department of Mechanical Engineering and Aeronautics, University of Patras
Rio, Patras, Greece
alexokos@lms.mech.upatras.gr

2024 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC) | 979-8-3503-6243-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICE/ITMC61926.2024.10794347

***Abstract*—The integration of human knowledge and experience with artificial intelligence, especially in the context of Industry5.0, holds the promise of advanced capabilities for manufacturing that may facilitate reduced waste and increased efficiency. However, there is a gap between the two. This work discusses the critical role of Explainable AI (XAI) within this paradigm, fostering a collaborative environment where human operators can leverage AI-driven insights. A framework for data-driven proactive quality control is coupled with XAI and human-centric approaches to enable a path towards zero-defect manufacturing processes, improved operational efficiency, and enhanced workforce empowerment. Furthermore, practical implications, the impact of XAI and recommendations for upskilling and reskilling the manufacturing personnel are discussed with a focus on small and medium-sized enterprises.**

***Keywords*—Explainable AI, Industry 5.0, Proactive Quality Control, Manufacturing**

I. Introduction

Quality control is an integral part of a production process [1], aiming at identifying defects, at different stages of production; with traditional quality control approaches including testing and visual or manual inspection [2]. Manufacturing systems, typically involve quality control in their manufacturing processes, where randomly selected products or a batch of products or even every product is inspected for defects [3]. Manual quality control techniques typically rely on operators' expertise and are time-consuming, cost-ineffective and prone to errors [4]. Such techniques are designed to leverage operators' expertise. This makes them hard to replicate while little explainability or insights regarding the root cause of the generation of defects can be extracted [5].

In the shift towards Industry 5.0, quality control integrates human insights with AI-driven automation to enhance the adaptability and the personalization of manufacturing processes but also to make them more proactive [6]. Unlike Industry 4.0, where the focus was primarily on automation and predictive analytics to preemptively identify defects [7][8][9], Industry 5.0 emphasizes a symbiotic interaction between human operators and technology [10]. This approach allows for real-time adjustments in production lines, based on both data-driven insights and human experience, effectively reducing waste and improving product quality and work satisfaction, putting the human at the centre of the production [11][12].

Nevertheless, modern AI-driven solutions lack the capability of reasoning about their outcomes and recommendations [13]. This is due to the nature of complex algorithms that are considered "black boxes" because their internal workings and decision processes are not transparent [14][15][16]. In this context, Explainable AI (XAI) can be used to make AI decisions transparent and understandable to human operators, thereby not only increasing trust but also enabling operators to make informed decisions about when and how to intervene in the production process [17]. Thus, the integration of human insight with AI in Industry 5.0 allows for dynamic and responsive quality control systems.

This paper explores a framework that integrates a digital twin model, predictive analytics for defect detection, and shapley additive explanations (SHAP) [18] XAI to enhance the collaboration between human operators and automated AI systems within Industry 5.0. We detail how these technologies converge to provide a holistic view of predictive analytics, empowering operators with a deeper understanding of AI decision-making and the implications of their actions. Moreover, this study discusses the broader implications of such integrations for workforce development and policymaking, with a special focus on the practical deployment of these advanced technologies in small and medium-sized enterprises.

II. Literature Review

XAI techniques can be classified into two distinct categories which include transparent and post-hoc models [19]. The transparent method is applied in AI techniques whose inner workings and decision-making processes are simple to interpret and represent, while post-hoc is applied in AI techniques built with data that are characterised by non-linear relationships or high data complexity [19]. A transparent model is characterised by its high simulatability, decomposability and algorithmic transparency; with models like linear/logistic regression, decision trees, K-nearest neighbours, and rule-based models some AI transparent models [20]. Lastly, post-hoc methods are grouped into two categories; including the model-agnostic and model-specific categories, as described in [19].

Model-specific XAI techniques support explainability constraints targeting an AI Deep Learning algorithm and its internal structure [21]. Model-specific techniques include feature relevance, example-based explanation, rule-based learning, and feature-based saliency maps [21].

In [22] a systematic review is performed on the use of feature relevance in anomaly detection algorithms, concluding that while approaches such as perturbation-based approaches, and gradient-based approaches can pinpoint important features of the models, they can also potentially introduce biases; such as in the case of gradient-based approaches may highlight features with high numerical gradients but ignore others, equally important, that don't exhibit strong gradient signals. Furthermore, in [23] the use of example-based explanations is discussed, whose goal is the clarification of why a specific decision was made as opposed to another potential decision by an algorithm. As discussed in [23], example-based explanations use historical data to evaluate the model's correctness, offering users insight into the application of advice across different scenarios. However, despite an increase in the persuasiveness of the advice, example-based explanations may not enhance a user's understanding of the model's outputs. Rule-based learning is examined in [24], where the ability to enhance transparency and understandability of decisions in complex AI systems is documented. However, rule-based learning is not suited to dynamic and unpredictable environments due to the extensiveness and complexity of rule sets often accompanying such use cases. Lastly, in [25] feature-based saliency maps are discussed, where pixels were inserted and removed from images used to train a DL algorithm, in a computer vision application, to determine the importance of different pixels and visualise the calculated importance in heat maps. Nevertheless, this technique is computationally heavy and requires constant optimisation of mask parameters [25].

Model-agnostic XAI techniques typically analyse a model's inputs and outputs, aiming at interpreting the model's behaviour to generate explanations [26]. Model-agnostic techniques include the local interpretable model-agnostic explanations (LIME), the shapley addictive explanations (SHAP), and layer-wise relevance propagation (LRP) [19],[26].

LIME is discussed in [27], where its strengths and limitations are reviewed. LIME is perturbing the input data to observe how the model's outputs change; thus understanding features with significant influence in the output. LIME provides explanations for individual outputs which increases its ability to understand the underlying decision-making process of models at a local level. However, LIME can potentially provide misleading interpretations due to its focus on local explanations rather than providing a global overview of the model's behaviour [28]. SHAP was introduced in [18]. SHAP can be used for both local and global explanations; increasing its versatility and ability to explain the global model's behaviour [18]. Nevertheless, SHAP's effectiveness in explaining a model's results is highly dependent on the sample of the dataset used by SHAP to build the explanation model and SHAP can be computationally intensive; however, this is a characteristic of all model-agnostic techniques [29]. Lastly, LRP is examined in [30]. LRP shifts its focus towards the output layer and progresses in a backwards manner to the input layer, proportionally distributing relevance scores to each neuron based on their contribution to the final output [30]. Nevertheless, LRP is accompanied by potential misinterpretations since the interpretations provided by LRP heavily rely on the underlying model's accuracy and correct implementation of the propagation rules [31].

Aiming to assist the human workforce, manufacturers are increasingly adopting XAI techniques together with their existing AI solutions to facilitate the efficient decision-making of personnel to improve quality control [32]. In [33] a system utilising XAI, based on LIME, provides insights on predictions generated by AI algorithms regarding defective rotating machines. Providing clear explanations of the predictions generated by the AI models, enhanced decision-making is achieved by providing equipment maintenance recommendations to ensure optimal equipment performance to prevent the generation of defects [33]. Additionally, in [34] XAI was used to improve the classification of defects in aircraft components, while in [35] the importance of human-centric approaches in quality control is emphasized, pinpointing that XAI can be an enabling technology to achieve operator support in an automated quality control environment.

In conclusion, the literature review underscores a significant demand for increased insight into the modern quality control approach employed in the manufacturing industry. Existing strategies often focus on adopting XAI to indirectly influence quality control in manufacturing, such as through predictive maintenance. In this study, the focus relies on describing the behaviour of the entire quality control system using an XAI layer running on top of predictive AI analytics and a digital twin providing recommendations to operators to enhance proactive quality control.

III. Methodology

This study aims to introduce a framework that couples XAI and a human-centric layer with advanced digital solutions, powered by complex AI models, aiming at increasing the understanding of results generated by AI models and incorporating a feedback loop to improve the AI models' performance and XAI model's explainability. An overview of the proposed methodology is illustrated in Fig. 1.

The framework's core is a digital twin. The digital twin is the digital representation of the physical manufacturing system providing the system's behaviour using AI-enabled approaches [36].

Apart from the behaviour of the manufacturing processes in the system, the digital twin also provides the behaviour of the product as it traverses the different processes. In conjunction with the digital twin, AI-powered data analysis modules access the behaviour of the product and processes as well as real-time data from the physical manufacturing system. Enabling proactive quality control, the AI data analysis modules utilise the provided product and process behaviour and real-time production data to predict future product defects.

Positioned on top of the digital twin and AI-enabled quality assessment models, an XAI layer provides explanations generated by the digital twin and AI models to human workers. The formation of the XAI layer is performed in three sequential steps, which include:

- ***Step 1 - Access product/process behaviour & AI model predictions***: The digital twin and AI-enabled predictive model feed behavioural and prediction data to the XAI layer, using dedicated components facilitating data access,

- ***Step 2 - SHAP values calculation***: Using the provided data by the digital twin and AI-enabled predictive model, SHAP values are dynamically calculated for each set of data points,
- ***Step 3 - User presentation***: Explanations generated by SHAP are presented to the human worker, through a User Interface of the system.

In greater detail, to guarantee a robust and constant feed of data from the digital twin and AI-enabled data analytics modules, a dedicated data access layer sits between the XAI layer and the rest of the system. Once the data is accessed cooperative game theory algorithms are employed to calculate the SHAP values. The algorithms measure each data point's contribution to the behaviour of the product and process as well as to the contribution to the predictive model's output. The SHAP values are calculated using (1) [18], and are dimensionless.

$$\varphi_i(f,x) = \sum_{z' \subseteq x'} \frac{|z'|(M-|z'|-1)}{M!}[f_x(z') - f_x(\frac{z'}{i})] \quad (1)$$

Where:

- φ: the impact of the data point i for the AI model f at a record x,
- $|z'|$: the number of data points in the subset z',
- M: the total number of data points,
- $f_x(z')$: the prediction of the AI model when the data points in the subset z' are used along with the data point i,
- $f_x(\frac{z'}{i})$: the prediction of the AI model when the data points in the subset z' are used without the data point i.

Together with the XAI layer, a human-centric interaction layer collects and processes feedback from human workers regarding the explanations of the XAI layer and the predictions of the AI model. The formation of the human-centric interaction layer is performed in four sequential steps, which include:

- ***Step 1 - Collect human worker feedback***: Through the user interface, the human worker can provide feedback on both the AI model's results and the explanations provided by the XAI layer,
- ***Step 2 - Feedback categorisation***: The provided feedback is categorised to facilitate its consumption by the human-centric interaction layer,
- ***Step 3 - Feedback analysis***: Based on the categorisation of the feedback, it is processed to identify repetitive feedback patterns,
- ***Step 4* - *Corrective actions application*:** Based on the feedback analysis results, corrective actions are taken with corrections applied to the XAI layer and AI models.

Delving deeper into the human-centric interaction layer, the user interface prompts the user to provide feedback on the results provided by the digital twin, AI models and the XAI

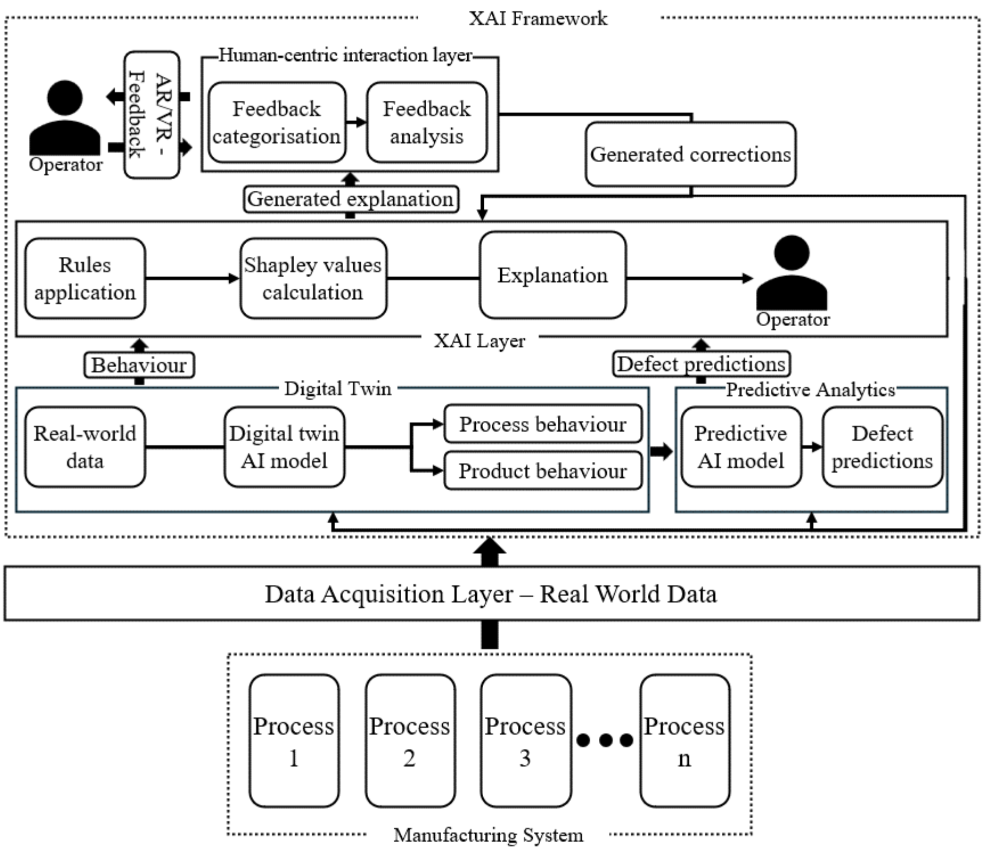

Fig. 1. The XAI Framework.

layer. Upon feedback receival, it is categorised against predetermined categories. The categories include, "feature importance correction", "defect prediction correction", "digital twin behaviour correction", "explanation misunderstood", "explanations helpful", and "predictions accurate". Additionally, the specific data points used during the AI models' execution and their output are grouped with the provided feedback to facilitate the feedback analysis and generation of corrective actions.

Following the feedback categorisation, the feedback is systematically analysed to determine common patterns within feedback data and corrective actions are applied to the AI models and XAI layer. To analyse feedback data statistical analysis approaches are pursued. A combination of association rule mining and correlation analysis is employed to uncover a high proportion of feedback about specific categories, with correlation analysis pinpointing the data points influencing most of the collected feedback. Ultimately, based on the feedback analysis alterations are performed in the AI models to increase their performance as well as alterations on the XAI layer to improve its explanation generation approach; reducing confusion caused by the explanation visualisation approach employed.

Lastly, the XAI framework includes a visualisation environment utilising AR/VR technology to enable humans to interact with the system to get a better understanding of the manufacturing systems operations and the areas where AI solutions are applied along with their outputs, presented interactively, to enhance their AI literacy and familiarity with the AI and XAI concept. Lastly, through AR/VR technology, new human workers in the production line can become easily accustomed and familiar with the entire manufacturing system, given the increased immersity such technologies offer compared to traditional desktop or mobile interfaces, displaying manufacturing assets using a 2D/3D representation.

IV. Use Case

The XAI framework has been applied to a virtual manufacturing environment to test its performance and suitability in assisting humans in understanding the outputs of AI models. The virtual manufacturing environment replicates a real-world robotic welding station, part of a larger manufacturing system capable of producing products for the automotive industry. The virtual robot welding process is of key importance due to its potential high welding defect generation and the important role of operators in proactively reducing defects, assisted by advanced digital solutions.

The virtual environment's backbone is a digital twin. The digital twin is powered by an AI model responsible for providing real-time high-quality synthetic data of the welding process including the welding speed, the peak welding temperature, the current, and the voltage. The real-time data collected by the digital twin are used by a predictive AI model that predicts potential welding defects. Welding defects include spots on the weld due to bad welding, laser cuts on the product due to the laser welding process and missing welding. The labelling of the synthetically generated data was performed in conjunction with human experts in the welding process, and *0* indicates defect absence, and *1* presence.

Humans interact with the virtual manufacturing system through a user interface that displays the results of predictive analytics, and based on the results of the AI predictive model, corrective actions on the process need to be performed by the operator. Assisting in the decision-making process the SHAP XAI's explanations are displayed to the operator alongside the predictive model's outputs.

The virtual environment was constructed in a development environment consisting of a Windows PC equipped with an Intel Core i7-13700H processor, 32 GB of DDR5 RAM, and a cuda-enabled NVIDIA RTX A1000 GPU with 6 GB GDDR6 VRAM, running Windows 11 Pro version 23H2. Python was used in the creation of the AI models and SHAP, together with its libraries such as PyTorch for the model creation. The digital twin was constructed using Node-RED and the React.js framework.

The application of SHAP necessitates a deep learning model capable of generating accurate defect results. The model used in the virtual environment is a custom deep-learning model. The model's architecture can be seen in Fig. 2. The AI model utilises as input the real-time synthetically generated data of the digital twin, and the predictions generated are then provided to the XAI layer.

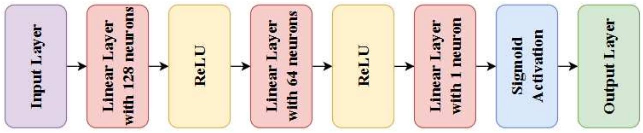

Fig. 2. Architecture Of The Predictive AI Model.

The predictive AI model was trained with 77,120 data points and its performance was validated using 19,280 data points. The data points represent the total number of values included in the dataset used to train and validate the predictive AI model. The dataset was composed based on the data generated by the digital twin and a representation of the dataset can be seen in Table I. Lastly, the dataset includes a label against welding defects that the model aims to predict.

TABLE I. Sample Of The Dataset Used To Train & Validate The Predictive Model.

Dataset Sample				
Welding speed (mm/min)	***Peak Welding Temperature (ºC)***	***Current (A)***	***Voltage (V)***	***Label***
127.2	1587.8	25.26	21.26	0
133.2	1588.2	72.51	17.93	1

The model was capable of predicting the presence or absence of defects given a set of real-time data provided by the digital twin. Given the nature of the task (predicting the presence or absence of welding defects), the model performs binary classification. The model's performance was measured using the validation data and the following performance metrics (2), (3), (4), (5).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{2}$$

$$Precision = \frac{TP}{TP + FP}, \tag{3}$$

$$Recall = \frac{TP}{TP+FN}, \quad (4)$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \quad (5)$$

Where:

- *TP is* the true positives,
- *FP* is the false positives,
- *FN* is the false negatives, and
- *TN* is the true negatives of the classification.

The deep learning model's performance can be seen in Table II.

TABLE II. PREDICTIVE MODEL'S PERFORMANCE.

Initial Predictive Model's Performance	
Performance metric	***Values***
Accuracy	0.85
Precision	0.84
Recall	0.84
F1-score	0.84

Given the relatively high performance of the deep learning algorithm, SHAP was applied. High performance of the AI model is essential to the correct application of SHAP to ensure accurate explanations and avoid misinterpretations, as discussed in section II. In the SHAP application, the *GradientExplainer* function was used. The *GradientExplainer* function was selected due to its high efficiency compared to alternative explanation functions, its ability to utilise the internal structure of the deep learning predictive model; resulting in more accurate and theoretically sound explanations, and its suitability in explaining neural network algorithms. In essence, the *GradientExplainer* function approximates the SHAP values provided by (1), due to the computational complexity a direct calculation of (1) would introduce. The performance of SHAP was measured using local accuracy, and additivity, given by (6), and (7), and can be seen in Table III.

$$f(x) = \varphi_0 + \sum_{i=1}^{M} \varphi_i(x), \quad (6)$$

$$f(X) = \varphi_0 + \sum_{i=1}^{M} \varphi_i(X), \quad (7)$$

Where:

- *f(x)*: the prediction for a specific instance *x*.
- φ_0: the expected value of the model outputs over the background dataset using the SHAP calculation,
- $\varphi_i(x)$: the SHAP values for each data point *i* for the instance *x*,
- *M*: the total number of data points,
- *f(X)*: the predictions across all instances in a set *X*.

The SHAP's performance can be seen in Table III.

TABLE III. THE PERFORMANCE OF SHAP.

Performance of SHAP	
Performance metric	***Values***
Local accuracy	0.2
Additivity	0.4

The results of local accuracy that are presented in Table III indicate that the SHAP has a moderate degree of error when SHAP values combined with the model's expected values, approximate an individual prediction; suggesting that the SHAP requires further adjustments to improve its explanations. Additionally, the additivity score of 0.4 from Table III points out that the overall discrepancy between the summed contributions of the SHAP values across all predictions and the actual outputs of the model is moderate; indicating that the SHAP can provide some valuable insight into the model's behaviour. In general, it can be observed that while SHAP explanations can generally provide adequate approximations of the model's behaviour, they may struggle with more specific predictions where the model's decision-making process is more complex.

Provided the relatively adequate ability of SHAP to explain the predictive AI model's behaviour and the importance of each feature in reaching the results shown to the user, summary plots are generated by SHAP to be displayed to the user. The SHAP plots explain to the human worker the

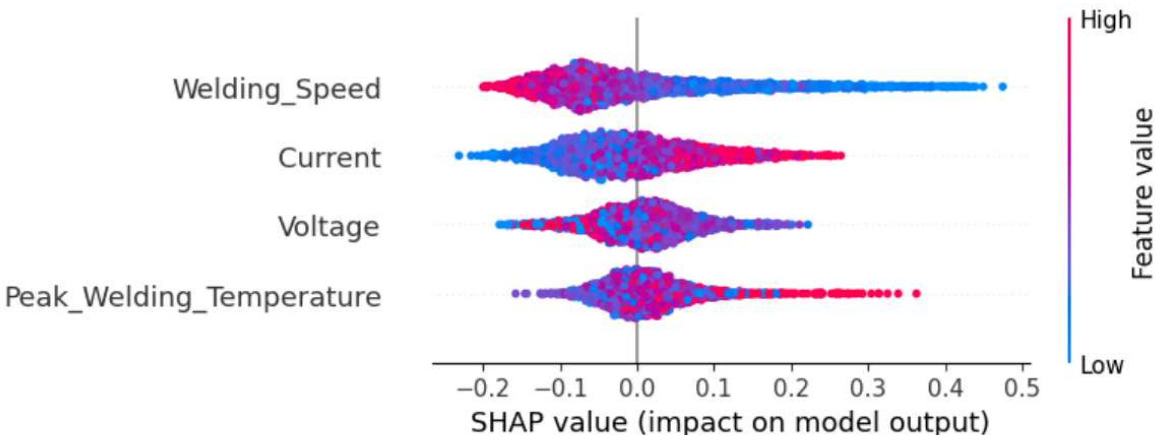

Fig. 3. The Summary Plot Of the SHAP For The Importance Of Features In Predicting Welding Defects.

importance of each feature used during the AI model's training in reaching a decision; which is the prediction of the presence or absence of a welding defect. The summary plot was configured to provide the importance of each feature against the prediction of a welding defect, and it can be seen in Fig. 3. In the summary plot, the features are displayed on the vertical axis. The SHAP values are indicated in the horizontal axis. The colour represents the importance of each data point of each feature, and points to the right from the *0* SHAP value increase the likelihood of the positive class; in this case the presence of welding defects. For example, for the "Current", the significant presence of high values on the positive side of the SHAP values points to the "Current" feature being of high significance in predicting the positive class; which in this case is the presence of welding defects.

An experiment was formulated, where four humans interacted with the virtual manufacturing environment. Humans were exposed to the results of the predictive AI model as well as the explanations generated by the XAI layer. Using the XAI framework the humans included in the experiment provided recommendations on the defects prediction model based on the importance and explanation generated for each data point, provided by SHAP. The feedback was used to adjust the AI model and its performance was recalculated as seen in Table IV.

TABLE IV. PREDICTIVE MODEL'S PERFORMANCE AFTER CORRECTIONS.

Predictive Model's Performance After Corrections		
Performance metric	***Values***	**Recorded difference against initial model**
Accuracy	0.89	4.7%
Precision	0.88	4.7%
Recall	0.87	3.6%
F1-score	0.87	3.6%

The results of Table IV indicate that through the XAI framework, the predictive AI model's performance was improved by an average of 4.15% across the four performance metrics. Nevertheless, it is worth pointing out here, that two of the four humans included in the experiment had prior experience with AI models, which could have played an important role in providing accurate feedback to the human-centric interaction layer of the framework. However, results may differ in an actual manufacturing environment where human operators, who interact with the framework, may have limited to no prior exposure to AI systems.

V. DISCUSSION

Effective integration of XAI in Industry 5.0 begins with the strategic selection of appropriate XAI models, such as SHAP, tailored to meet specific operational needs for detailed feature attribution. This selection process must ensure that the chosen models align well with existing operational workflows. Organisations are encouraged to adopt an iterative approach to enhance XAI efficacy, initiating pilot projects that allow for continuous feedback-driven refinements. This method not only builds familiarity with XAI systems but also strengthens trust and confidence among operators.

In conjunction with technical integration, the development of robust training programs is crucial. These programs should focus on enhancing AI literacy, demonstrating practical applications of XAI, and providing hands-on experience with AI systems. By educating workers about both the functionalities and limitations of AI and specifically illustrating how tools like SHAP deliver interpretable insights, the workforce can more effectively understand and leverage AI outputs.

Furthermore, the successful adoption and utility of XAI require the support of thoughtful policy frameworks that incentivise innovation while upholding ethical standards. Such policies should foster transparency and accountability, ensuring that AI and XAI systems are both auditable and their decisions contestable by human operators.

Additionally, promoting cross-sector collaboration through forums and consortia is essential to standardise XAI practices across industries. This approach not only simplifies the adoption process, particularly for small and medium-sized enterprises but also facilitates the dissemination of best practices and experiences. Consequently, it accelerates the practical deployment of XAI and ensures its benefits are broadly accessible.

Lastly, in an actual manufacturing environment, which highly relies on operators' expertise in the identification of welding-related defects, an accepted average defect detection rate is approximately 85%. This signals that the original AI predictive model can not be considered up to par with manual inspection techniques. Thus, an average improvement of 4% (as seen in Table IV) of a predictive AI system assisting in the early identification of defects could be of high importance. Such an improvement would directly impact the total number of defects that can be identified at an early stage of production, thus advancing the quality control process. Nonetheless, further improvement of the approach, targeting the improvement of SHAP could be critical in optimising the proposed XAI framework. Lastly, while a learning curve exists to use the framework, the high potential for improving AI systems based on human experts' feedback alleviates the potential man-hours and associated costs required to train new teams to utilise such tools in day-to-day operations.

VI. CONCLUSION

This paper discussed the integration of XAI within Industry 5.0, emphasizing its critical role in bridging the gap between advanced automation technologies and human expertise. By implementing a framework that combines digital twins, predictive analytics, and SHAP-based XAI, this study supported that such integration can enhance proactive quality control, enabling a move towards zero-defect manufacturing processes.

The potential of XAI to make AI systems more transparent and understandable has proven substantial, empowering human operators to trust and effectively interact with AI-driven systems. This transparency fosters a more inclusive work environment, where technology augments rather than replaces human skills, potentially yielding broader economic and social benefits and promoting a collaborative, and adaptive workforce.

Nevertheless, the proposed methodology presents complexities and high computational needs, particularly with real-time SHAP implementation. Additionally, integrating sophisticated XAI models into legacy systems poses significant challenges.

Future research will focus on optimising XAI algorithm efficiency for real-world applications and broadening the adaptability of XAI systems to accommodate diverse manufacturing scenarios and operator expertise levels.

ACKNOWLEDGEMENT

This work was partially supported by the HORIZON-CL4-2021-TWIN-TRANSITION-01 openZDM project, under Grant Agreement No. 1010586.

REFERENCES

[1] M. Colledani et al., “Design and management of manufacturing systems for production quality,” CIRP Annals, vol. 63, no. 2, pp. 773–796, 2014.

[2] J. Butt, R. Bhaskar, and V. Mohaghegh, “Non-Destructive and Destructive Testing to Analyse the Effects of Processing Parameters on the Tensile and Flexural Properties of FFF-Printed Graphene-Enhanced PLA,” J. Compos. Sci., vol. 6, no. 5, p. 148, May 2022.

[3] G. Chryssolouris, Manufacturing systems: theory and practice, 2nd ed. in Mechanical engineering series. New York: Springer, 2006.

[4] V. Medici et al., “Integration of Non-Destructive Inspection (NDI) systems for Zero-Defect Manufacturing in the Industry 4.0 era,” in 2023 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT), Brescia, Italy: IEEE, Jun. 2023, pp. 439–444.

[5] Y. Torres, S. Nadeau, and K. Landau, “Classification and Quantification of Human Error in Manufacturing: A Case Study in Complex Manual Assembly,” Applied Sciences, vol. 11, no. 2, p. 749, Jan. 2021.

[6] M. H. Zafar, E. F. Langås, and F. Sanfilippo, “Exploring the synergies between collaborative robotics, digital twins, augmentation, and industry 5.0 for smart manufacturing: A state-of-the-art review,” Robotics and Computer-Integrated Manufacturing, vol. 89, p. 102769, Oct. 2024.

[7] G. Dutta, R. Kumar, R. Sindhwani, and R. Kr. Singh, “Digitalization priorities of quality control processes for SMEs: a conceptual study in perspective of Industry 4.0 adoption,” J Intell Manuf, vol. 32, no. 6, pp. 1679–1698, Aug. 2021.

[8] P. Catti, N. Nikolakis, K. Sipsas, N. Picco, and K. Alexopoulos, “A hybrid digital twin approach for proactive quality control in manufacturing,” Procedia Computer Science, vol. 232, pp. 3083–3091, Jan. 2024.

[9] N. Leberruyer, J. Bruch, M. Ahlskog, and S. Afshar, “Toward Zero Defect Manufacturing with the support of Artificial Intelligence—Insights from an industrial application,” Computers in Industry, vol. 147, p. 103877, May 2023.

[10] X. Xu, Y. Lu, B. Vogel-Heuser, and L. Wang, “Industry 4.0 and Industry 5.0—Inception, conception and perception,” Journal of Manufacturing Systems, vol. 61, pp. 530–535, Oct. 2021.

[11] A. Hanif et al., “A Comprehensive Survey of Explainable Artificial Intelligence (XAI) Methods: Exploring Transparency and Interpretability,” in Web Information Systems Engineering – WISE 2023, vol. 14306, F. Zhang, H. Wang, M. Barhamgi, L. Chen, and R. Zhou, Eds., in Lecture Notes in Computer Science, vol. 14306. , Singapore: Springer Nature Singapore, 2023, pp. 915–925.

[12] N. V. Naveen Vemuri, “Enhancing Human-Robot Collaboration in Industry 4.0 with AI-driven HRI,” pst, vol. 47, no. 4, pp. 341–358, Dec. 2023.

[13] J. Maclure, “AI, Explainability and Public Reason: The Argument from the Limitations of the Human Mind,” Minds & Machines, vol. 31, no. 3, pp. 421–438, Sep. 2021.

[14] P. N. Mueller, L. Woelfl, and S. Can, “Bridging the gap between AI and the industry — A study on bearing fault detection in PMSM-driven systems using CNN and inverter measurement,” Engineering Applications of Artificial Intelligence, vol. 126, p. 106834, Nov. 2023.

[15] W. Gao et al., “Machine tool calibration: Measurement, modeling, and compensation of machine tool errors,” International Journal of Machine Tools and Manufacture, vol. 187, p. 104017, Apr. 2023.

[16] C. Ferreira and G. Gonçalves, “Remaining Useful Life prediction and challenges: A literature review on the use of Machine Learning Methods,” Journal of Manufacturing Systems, vol. 63, pp. 550–562, Apr. 2022.

[17] S. Huang, B. Wang, X. Li, P. Zheng, D. Mourtzis, and L. Wang, “Industry 5.0 and Society 5.0—Comparison, complementation and co-evolution,” Journal of Manufacturing Systems, vol. 64, pp. 424–428, Jul. 2022.

[18] S. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions.” arXiv, Nov. 24, 2017.

[19] P. Gohel, P. Singh, and M. Mohanty, “Explainable AI: current status and future directions.” arXiv, Jul. 12, 2021.

[20] T. P. Quinn, S. Gupta, S. Venkatesh, and V. Le, “A Field Guide to Scientific XAI: Transparent and Interpretable Deep Learning for Bioinformatics Research.” arXiv, Oct. 13, 2021.

[21] W. Saeed and C. Omlin, “Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities,” Knowledge-Based Systems, vol. 263, p. 110273, Mar. 2023.

[22] J. Tritscher, A. Krause, and A. Hotho, “Feature relevance XAI in anomaly detection: Reviewing approaches and challenges,” Front. Artif. Intell., vol. 6, p. 1099521, Feb. 2023.

[23] J. Van Der Waa, E. Nieuwburg, A. Cremers, and M. Neerincx, “Evaluating XAI: A comparison of rule-based and example-based explanations,” Artificial Intelligence, vol. 291, p. 103404, Feb. 2021.

[24] B. M. Keneni et al., “Evolving Rule-Based Explainable Artificial Intelligence for Unmanned Aerial Vehicles,” IEEE Access, vol. 7, pp. 17001–17016, 2019.

[25] Z. Wang and I. Joe, “OISE: Optimized Input Sampling Explanation with a Saliency Map Based on the Black-Box Model,” Applied Sciences, vol. 13, no. 10, p. 5886, May 2023.

[26] L. Gianfagna and A. Di Cecco, “Model-Agnostic Methods for XAI,” in Explainable AI with Python, Cham: Springer International Publishing, 2021, pp. 81–113.

[27] J. Dieber and S. Kirrane, “Why model why? Assessing the strengths and limitations of LIME.” arXiv, Nov. 30, 2020.

[28] E. Amparore, A. Perotti, and P. Bajardi, “To trust or not to trust an explanation: using LEAF to evaluate local linear XAI methods,” PeerJ Computer Science, vol. 7, p. e479, Apr. 2021.

[29] K. Roshan and A. Zafar, “Utilizing XAI Technique to Improve Autoencoder based Model for Computer Network Anomaly Detection with Shapley Additive Explanation(SHAP),” IJCNC, vol. 13, no. 6, pp. 109–128, Sep. 2021.

[30] I. Ullah, A. Rios, V. Gala, and S. Mckeever, “Explaining Deep Learning Models for Tabular Data Using Layer-Wise Relevance Propagation,” Applied Sciences, vol. 12, no. 1, p. 136, Dec. 2021.

[31] S. G. Kim, S. Ryu, H. Kim, K. Jin, and J. Cho, “Enhancing the Explainability of AI Models in Nuclear Power Plants with Layer-wise Relevance Propagation,” presented at the Proceedings of the Transactions of the Korean Nuclear Society Virtual Autumn Meeting, Jeju, Korea, 2021, pp. 21–22.

[32] Z. Alexander, D. H. Chau, and C. Saldaña, “An Interrogative Survey of Explainable AI in Manufacturing,” IEEE Trans. Ind. Inf., pp. 1–13, 2024.

[33] S. Gawde, S. Patil, S. Kumar, P. Kamat, and K. Kotecha, “An explainable predictive maintenance strategy for multi-fault diagnosis of rotating machines using multi-sensor data fusion,” Decision Analytics Journal, vol. 10, p. 100425, Mar. 2024.

[34] S. Meister, M. Wermes, J. Stüve, and R. M. Groves, “Investigations on Explainable Artificial Intelligence methods for the deep learning classification of fibre layup defect in the automated composite manufacturing,” Composites Part B: Engineering, vol. 224, p. 109160, Nov. 2021.

[35] C. Bechinie, S. Zafari, L. Kroeninger, J. Puthenkalam, and M. Tscheligi, “Toward human-centered intelligent assistance system in manufacturing: challenges and potentials for operator 5.0,” Procedia Computer Science, vol. 232, pp. 1584–1596, 2024.

[36] M. Grieves, “Origins of the Digital Twin Concept,” 2016.