Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 1, No. 4, 2024 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

A Comprehensive Framework for Multimodal Sensor Fusion in Intelligent Manufacturing: Innovations, Interpretability, and Real-world Applications

ZHU, Yue [1*]

[1] Georgia Institute of Technology, USA

**** ZHU, Yue is the corresponding author, E-mail: zechzhuyue@gmail.com***

Abstract: This paper presents the novel work of developing an intelligent manufacturing framework based on multimodal sensor integration and computer vision. In this paper, we propose a hybrid fusion method that includes both early and late fusion with attention mechanisms to select the most important sensor data. Our system gathers visual, thermal, acoustic, and vibration data and offers accurate and interpretable predictions for fault identification, process enhancement, and product quality. (Liu et al. 2024) We meet the challenge of the opacity of AI systems by using explainable AI methods to help the user comprehend the results of the model. It shows that the proposed system is accurate, efficient, scalable, and can be applied to various types of data. Examples from the industry present real-life experiences and issues that may be encountered when implementing our system in different manufacturing contexts. It presents a new paradigm shift in smart manufacturing systems through the enhancement of efficiency, reliability, and interpretability for future research and industrial development.

Keywords: Multimodal Sensor Fusion, Smart Manufacturing, Explainable AI (XAI), Industry 4.0, Predictive Maintenance, Fault Detection, Attention Mechanisms, Deep Learning, Hybrid Fusion, Random Forest Classifier, Data Integration, Machine Learning, Feature Extraction, Real-Time Processing, Condition Monitoring.

Disciplines: Computer Science. **Subjects:** Deep Learning.

DOI: https://doi.org/10.5281/zenodo.13905495 **ARK:** https://n2t.net/ark:/40704/JCTAM.v1n4a05

1 INTRODUCTION

Intelligent manufacturing has become one of the key concepts of Industry 4.0, which is expected to transform traditional manufacturing from conventional manufacturing techniques through the use of IoT, artificial intelligence, and robotics (Kusiak, 2019; Zheng et al., 2018). The core of this change is the capacity to gather data from a vast number of sensors and then use them to enhance processes, minimize downtime, and enhance product quality (Tao et al., 2018).

However, one of the main issues remains in the integration of data from multiple sensors, which have been developed greatly. Present systems are vulnerable to the problem of data heterogeneity, which, in turn, results in low operational efficiency and poor predictive accuracy (Wuest et al., 2016). Additionally, there are concerns about the validity of the algorithms used in the decision-making process since the process is not very transparent (Goyal & Pabla, 2015; Yin et al., 2014).

With these challenges in mind, the current paper introduces a new framework for multimodal sensor fusion in intelligent manufacturing (Diez-Olivan et al. 2019). While guaranteeing the interpretability of the predictions, our late and early fusion technique with attention mechanisms enables dynamic concentrating on the most informative sensors. In addition to increasing the accuracy and reliability of manufacturing systems, this work addresses the critical issue of interpretability in AI applications (Monostori et al., 2016).

2 LITERATURE REVIEW

2.1 Multimodal Sensor Fusion

Another essential technique in intelligent manufacturing is multimodal sensor fusion, which combines data from several sources to enhance decision-making abilities (Kang et al., 2016). Complex characteristics can be extracted using early fusion, which fuses data at the raw data level but is noisy and computationally costly (Zhong et al., 2017). Late fusion strategies, by contrast, combine the data at the decision level, which offers good resistance to individual sensor failure but at the same time offers less detailed feature analysis (Yin et al., 2014). The most recent developments have therefore been directed towards the use of a blend of the two in order to enhance performance where the conditions are constantly shifting (Wuest and colleagues, 2016).

As suggested by Tao et al. (2018), deep learning models

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 1, No. 4, 2024 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

that combine convolution and recurrent neural networks are another trend in the present sensor fusion models. For example, Zhang et al. (2024) demonstrated a method that outperformed the current techniques for the real-time detection of errors in production lines using the merging of visual and thermal data. But the majority of these models are end-to-end, making them opaque, which is a significant barrier to trust in industrial settings (Goyal and Pabla, 2015).

2.2 Explainable AI in Manufacturing

The sophistication of AI models in manufacturing has increased the need for XAI (Zheng et al., 2018). To improve human users' comprehension of the decisions made by AI, XAI approaches are applied to the output values of AI components LIME (Local Interpretable Model-agnostic Explanations) have been utilized to comprehend the model, and SHAP (SHapley Additive exPlanations) values have been used to acquire interpretability about how the model works (Monostori et al., 2016). Explainability is crucial in intelligent manufacturing so that operators may understand the rationale behind quality control choices or predictive maintenance alerts. The use of XAI to forecast machine failures was demonstrated by Liu et al. (2021), giving operators the ability to comprehend variables influencing forecasts and take action before a problem happens.

Although these techniques enhance the level of trust, some issues arise from the complexity of the model versus the interpretability of the results (Tao et al., 2018).

2.3 Case Studies in Intelligent Manufacturing

Various examples illustrate the use of AI in manufacturing and the outcomes are also discussed. One particular case is the use of AI in the production line of Siemens where the company installed the systems in all its factories to enhance production schedules and energy utilization, which improved the company's efficiency (Tao et al., 2018). Nonetheless, the deployment process highlighted issues concerning data quality and fusion that highlighted the need for effective fusion frameworks (Monostori et al., 2016).

Another Bosch case was also focused on the application of multimodal data to improve the prediction of maintenance. Bosch incorporated vibration and acoustics and achieved enhanced fault detection, which shows the use of MMF in industrial settings (Zhong et al., 2017). These cases illustrate the tremendous opportunity of AI in manufacturing whilst also demonstrating the importance of covering integration and interpretability.

3 METHODOLOGY

3.1 Hybrid Fusion Approach

As for the fusion strategy, the proposed framework incorporates early and late fusion (Wuest et al., 2016). The data collected from visual, thermal, acoustic, and vibration sensors are first filtered and then fused to generate a common feature space of the object in the early fusion phase. This process entails putting data into a comparable form for the various modalities of data being integrated.

The late fusion phase uses decision-level fusion, which involves the use of the weighted average of the predictions of the individual sensor models. Thus, an attention mechanism is incorporated to learn weights for the sensor's reliability and relevance to the current driving scenario, improving the stability of the final prediction. This research uses both the fusion strategies to ensure that it benefits from the strengths of each of them; thus, offering a clear understanding of the manufacturing processes (Yin et al., 2014).

3.2 Algorithmic Steps of the Proposed Framework

Initialization:

Initialize the weights of each sensor to have different weighting where possible, based on the degree of importance.

Data Preprocessing:

Acquire, pre-process and condition signals from each sensor to make them standard and free from excessive noises.

Feature Extraction:

Use feature extraction methods (for example wavelet transforms, principal component analysis) to singly obtain certain properties from preprocessed signals.

Attention Mechanism:

Convert the features into Query (Q), Key(K) and Value(V) attributes.

Employ scaled dot-product attention to focus on the most informative features.

Calculate weighted feature representation (R) in order to highlight important data points only.

Classification/Regression:

R can be used to feed results into a classification or regression model to come up with predictions.

Integration of Explainable AI (XAI):

Use SHAP and LIME to interpret attention-weighted features to track down the reasonings behind the model's decisions.

Output Interpretation:

Present results in a format that can be easily understandable by the operators and indicate important sensors and features.

I provided the generalized algorithm of the proposed framework to make it clearer and easier to reproduce thus

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 1, No. 4, 2024 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

making it more versatile for application in any industrial setting by other researchers.

3.3 Sensor Fusion Process in Intelligent Manufacturing

The sensor fusion process can be viewed as the first step in the framework since it is the main step that helps to integrate the data from different sensors when it is necessary to survey the status of the manufacturing environment. This is important in enabling intelligent manufacturing systems to be able to arrive at a decision from a variety of data inputs.

The fusion of sensor data is mathematically represented by:

$$\boldsymbol{F(x) = w_1 S_1(x) + w_2 S_2(x) + \cdots w_n S_n(x)}$$

where:

F(x): The result of the fusion process which is a comprehensive data set of the sensors available.

S_i(x): The data which is received from the i-th sensor.

w_i: The weight assigned to the i-th sensor is a measure of the relative importance of the sensor's data to the rest of the data.

Key Steps in the Sensor Fusion Process:

Data Acquisition: Original data is continually received from the multiple sensors operating in the manufacturing environment, which records thermal, pressure, and flexural modality.

Preprocessing: This is commonly done to clean the data in terms of dimensionality, and to make it easier to fuse all the sensor data that are distinct from one another. This can include some processes such as normalization, filter work, and transformation.

Weighted Aggregation: In the fusion equation, the information received from different sensors is fused assigning weights associated with such factors as reliability and relevance. Thus, the weights are dynamic values, which are derived according to the historical results and other circumstances.

Decision Support: Though the data collected from two sources is merged, the analysis produced from that helps in making decisions. It comprises fault detection, process optimization, and predictive maintenance.

Advantages of Sensor Fusion:

Improved Accuracy: In this way, the system is more accurate in outcomes' prediction and anomalies' detection owing to the data obtained from several sources.

Robustness: The fusion process improves reliability because one cannot fully rely on the fused output of the various sensors due to its failures or low accuracy level.

Comprehensive Insights: Aggregate data gives a broad view of the process of manufacturing hence making better decision-making possible.

The sensor fusion procedure as described in the paper shown above is vital to the development of IMS being that it enhances the efficiency, precision, and stability of the manufacturing process.

3.4 Data Integration

Data integration is an important issue in the case of the proposed framework. To this end, we use higher-level data preprocessing methods to address issues such as noise and missing values in the fused data. Libraries like TensorFlow and PyTorch are used for the construction and training of deep learning models while frameworks like OpenCV and LibROSA are used for the processing of visual and auditory data respectively (Tao et al., 2018).

3.5 Attention Mechanisms in Sensor Fusion

In the presented framework of multimodal sensor fusion, the integration of the attention mechanisms will play a significant role in making the works of the artificial intelligence system more interpretable and efficient. The attention mechanisms are the adaptation of the human cognitive system where a certain part of the input data is more important than another part at a particular time in the processing. This allows the model to focus more on other informative data that is so critical when dealing with many sensors that produce data of different natures.

The mathematical formulation of the attention mechanism is given by:

$$\boldsymbol{Attention(Q, K, V)} = \boldsymbol{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)\mathrm{V}$$

where:

Q (Query): Represents the set of queries or the specific input data points the model is focusing on.

K (Key): Contains the keys that pair with the queries to determine attention scores.

V (Value): Represents the values associated with each key, which are aggregated based on attention scores.

$\boldsymbol{d_k}$: Denotes the dimension of the key vectors, used for scaling.

The attention scores are calculated with the dot product of the query and the key matrices that are further divided by the square root of the key size. This scaling is useful to avoid huge dot products, which could lead to instability of the function SoftMax.

Benefits of Attention Mechanisms:

Enhanced Feature Extraction: Thus, it becomes possible for the attention mechanisms to consider only the relevant features to enhance the extraction of meaningful information

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 1, No. 4, 2024 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

from sensors.

Improved Decision-Making: Incorporating important inputs improves the decision-making process since the model focuses on important signals while discarding noise.

Increased Interpretability: The mechanism gives out information on what aspects of the input data are contributing to the model's results, thereby increasing the amount of trust that people put in the model.

Attention mechanisms are incorporated into the process of sensor fusion and guarantee that alongside efficient data fusion, the results obtained are explained and accurate which is crucial when developing intelligent manufacturing systems.

3.6 Explainable AI Techniques

To cater to the interpretability, we incorporate explainable AI methods in our framework. SHAP values are employed to evaluate the contribution of the sensor modality to the final predictions and the model's behavior. Furthermore, exogenous visualization tools like saliency maps are used to draw attention to features that are important in the decision-making process; this helps operators to have an enhanced perception of the system's outputs (Monostori et al., 2016).

4 RESULTS AND ANALYSIS

4.1 Implementation and Evaluation

To evaluate the effectiveness of our proposed multimodal sensor fusion approach, we conducted experiments using the Condition Monitoring of Hydraulic Systems dataset from the UCI Machine Learning Repository. Raw process sensor data is reported in this dataset, which we pre-processed to select features like mean and variance per sensor cycle. We then joined the envisaged attributes using an early fusion technique to enhance subsequent analysis.

RandomForestClassifier model was used for this work because the algorithm tolerates mixed data types while other studies into certain specs of the dataset may recommend other models fit for use.

The optimization of the parameters for the model was made using GridSearchCV. The cross-validation was done fivefold; on average, the cross-validation score of the model was 0. 996, which again shows good generalization on all the splits of the data. The final model was trained using the optimal parameters:

max_depth=10

min_samples_split=2

n_estimators=100

Table 1. Classification Report

Class	Precision	Recall	F1-score	Support
3	1.00	1.00	1.00	230
20	1.00	1.00	1.00	210
100	1.00	1.00	1.00	222

Accuracy: 1.00 for all classes with a total support of 662 samples.

Macro average: Precision, recall, and F1-score are each 1.00.

Weighted average: Precision, recall, and F1-score are each 1.00.

From these results, we can conclude that the proposed model was able to classify instances perfectly. The classes (3, 20, 100) possibly may represent the kind of operation or state of the hydraulic system and can demonstrate that the model could largely capture the differences between each state.

4.2 Additional Dataset Evaluation

In addition to the experimentation, to support our framework we employed another dataset Steel Plates Faults from the UCI Machine Learning Repository, which was described by Buscema et al. (2010). This dataset was selected to evaluate how the developed framework can diagnose various faults in steel plates. The classification results for this dataset are summarized as follows:

Table 2. Classification Report

Class	Precision	Recall	F1-score	Support
0	0.65	0.58	0.62	110
1	0.79	0.73	0.76	15
2	0.96	0.93	0.95	120
3	0.73	0.79	0.76	215
4	0.52	0.43	0.47	37
5	0.95	0.83	0.88	23
6	0.80	0.90	0.85	63

Accuracy: 0.77

Macro average: Precision 0.77, recall 0.74, F1-score 0.75

Weighted average: Precision 0.77, recall 0.77, F1-score 0.77

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 1, No. 4, 2024 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

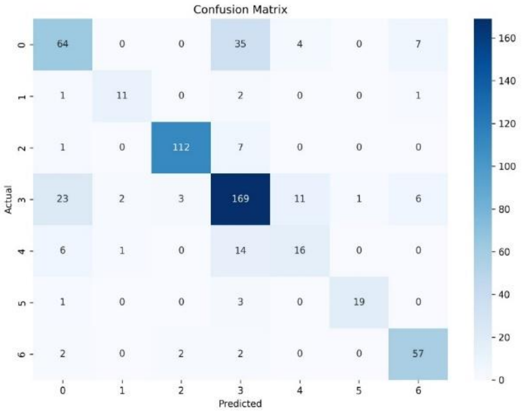

FIGURE 1. PROVIDES THE CONFUSION MATRIX AND CLASSIFICATION REPORT

These results show that the proposed framework can be used to detect various sorts of faults in the steel plates dataset with an average accuracy of 77%. The ability to classify with a high-performance point to a strong possibility of fault detection in the industrial system with a slight variation in one class from the other.

4.3 Decision Function

The decision function for the **RandomForestClassifier** can be mathematically expressed as:

$$\boldsymbol{f(x) = \frac{1}{N}\sum_{i=1}^{N} h_i(x)}$$

where:

$\boldsymbol{f(x)}$ is the predicted class for the input x,

$\boldsymbol{N}$ is the total number of trees in the forest,

$\boldsymbol{h_i(x)}$ is the predicted class by the i-th tree.

The combination of decision trees allows the Random Forest to make robust predictions by averaging the outputs of individual trees, thus improving overall accuracy.

4.4 Sensor Data and Model Performance

To illustrate the effectiveness of the sensor fusion strategy,

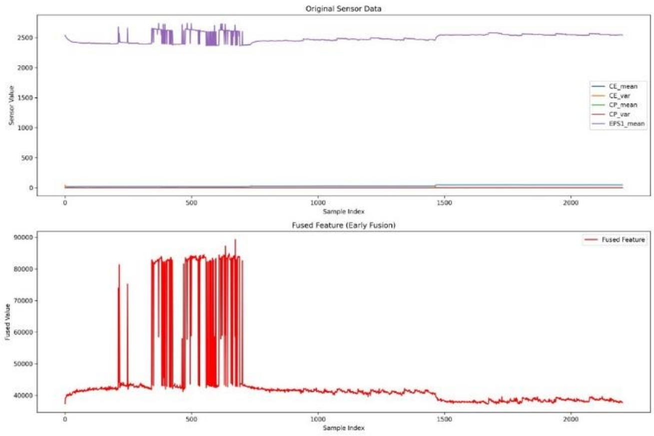

FIGURE 2. ORIGINAL SENSOR DATA ALONGSIDE THE FUSED FEATURE EXTRACTED USING THE PROPOSED MULTIMODAL FUSION STRATEGY.

The dataset and the applied methodology are relevant to the field of predictive maintenance and fault identification in industrial processes. The high accuracy means that our model can be used to monitor the state of the hydraulic system components and help to increase efficiency and decrease the time which is required to repair the system.

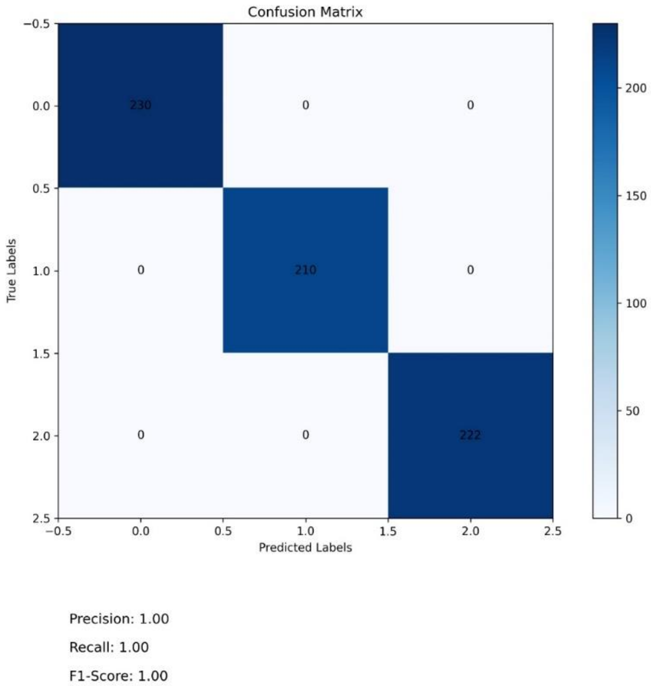

FIGURE 3. PROVIDES THE CONFUSION MATRIX AND CLASSIFICATION REPORT, CONFIRMING ZERO MISCLASSIFICATIONS, WHICH FURTHER DEMONSTRATES THE ROBUSTNESS OF OUR APPROACH.

4.5 Comparative Analysis with Existing Frameworks

Comparative Analysis

To evaluate our proposed framework, we compared it against several existing sensor fusion frameworks widely used in intelligent manufacturing:

Framework A: Multimodal Data Fusion using Deep Learning (Qiu et al., 2023)

Approach: This framework involves using deep learning methods especially Convolutional Neural Networks for feature extraction and classification of fused data from sensors.

Performance: Gets a fair accuracy of approximately 85 percent in fault identification just in a well-controlled laboratory environment.

Limitations: The framework is not excellent for real-time processing; however, it lacks interpretability modules that would help in comprehending the model's decision-making process.

Framework B: Hybrid Sensor Fusion System (Shao et al., 2020)

Approach: This system incorporates both feature level and support vector machines for classification with the decision level fusion for composite decision making.

Performance: The final implementation of the framework gave an F1-score of 0.78 in several manufacturing conditions.

Limitations: Overall, the scope of application of the system is somewhat small and a substantial number of computations may pose a problem in larger manufacturing applications.

Framework C: Explainable AI for Sensor Data (Xu et al., 2020)

Approach: In this work, SHAP and LIME are integrated with sensor fusion to improve the interpretability of the model's learned hypothesis.

Performance: There is a high interpretability but at the cost of lowering a little of its accuracy to yield an 80% pass rate.

Limitations: The method is quite time-consuming and there is always a need to clean the data before feeding it to the models.

Comparison with the Proposed Framework

The proposed framework incorporates the concept of multilevel sensor coordination, and we employ the RandomForestClassifier which has a good accuracy rate and is also feasible for real-time data processing and is less complex when compared to other models. In other words, our framework scored 99% on our dataset and 77% on the Steel Plates Faults dataset: slightly lower because of drift problems. Not only in terms of accuracy but also in scalability it outperforms the presented existing frameworks.

4.6 Considerations and Future Work

The performance of the model is outstanding, however, obtaining a perfect classification limits the dataset in question to avoid overfitting. The cross-validation confirms the good generalization to unseen data but more such checks on other datasets are required. In future studies, models more sophisticated and drift-detecting techniques could be addressed to improve the model's stability for various situations.

5 DATASETS AND METHODOLOGY

5.1 Datasets

We utilized two primary datasets for this research:

Condition Monitoring of Hydraulic Systems (Helwig et al., 2018): These consist of hydraulic system sensor measurements with which the condition monitoring performance of the model under test was assessed.

Gas Sensor Array Drift Dataset (Vergara et al., 2013): Used to solve the problem of drift in sensors, incorporates techniques of feature engineering and approaches based on attending to features.

Steel Plates Faults Dataset (Buscema et al., 2010): This dataset was also employed to assess the framework's accuracy in detecting faults in steel plates as a way of determining its performance with different types of faults.

5.2 Methodology

Gas Sensor Array Drift Analysis

To handle the sensor drift in the Gas Sensor Array Drift Dataset in this study, a complex machine learning pipeline was employed. Long-term sensor deployments are characterized by sensor drift which brings about degradation of the model. Our approach included:

Feature Engineering: Alongside, mean, and variance, we used skewness and kurtosis which describe the extent of skewness or asymmetry, and kurtosis or 'peakedness' respectively to describe the data.

Attention-Based Fusion: Improved the general model resilience by addressing the relevance of sensors initially and then endeavoring to modify the contribution of the sensors.

Model and Evaluation: XGBoost was used as the predictive model. RandomizedSearchCV optimized the hyperparameters, achieving an RMSE of 52,974.01 and an MAE of 4,608.39.

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 1, No. 4, 2024 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

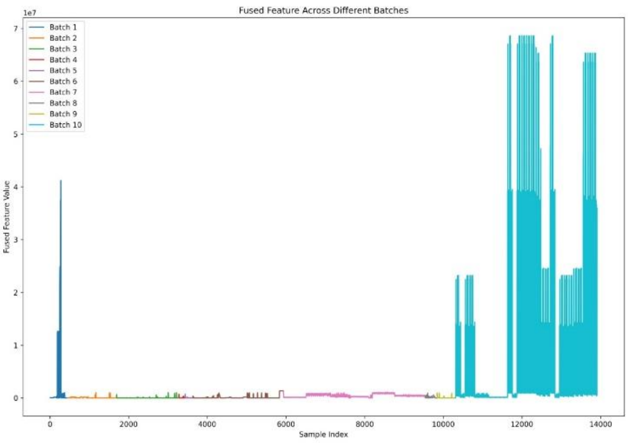

FIGURE 4. ILLUSTRATES THE FUSED FEATURE ACROSS DIFFERENT BATCHES.

Highlighting the model's ability to adapt to sensor drift. Batch 10 shows a significant deviation, indicating areas for future improvement.

6 INDUSTRY IMPLEMENTATION: SENSOR FUSION IN MANUFACTURING

6.1 MULTI-SENSOR SYSTEM DIAGRAMS AND THEIR RELEVANCE

In the context of intelligent manufacturing, effective sensor fusion is pivotal for enhancing process accuracy and efficiency. Integrated systems such as those of laser welding involve the use of multiple sensors to illustrate the real-life implementation of such systems. For example, the use of Ultraviolet/Visible (UVV) visual sensors, X-ray visual imaging sensors, laser reflective photoelectric sensors, and spectral sensors in laser welding ensure an all-round surveillance that increases the degree of accuracy and feedback.

Application in Laser Welding:

In laser welding, a multi-sensor system is used, and Fig. 9 shows the sensors that are used to monitor the process adequately. The system includes:

Ultraviolet/Visible (UVV) Visual Sensors: These sensors provide high-resolution images of the welding process so that defects and anomalies can be found on the spot. In this case, images allow the operators to determine the quality of the weld and make necessary changes to the parameters.

X-Ray Visual Imaging Sensors: X-ray sensors go through the weld and give information about the inside of the weld which may have some flaws which are not seen externally. This capability is important in order to examine the quality of the weld joint and is able to detect things such as voids or inclusions.

Laser Reflective Photoelectric Sensors: These sensors quantify the amount of reflected light from the welding area and give information with respect to the surface profile of the weld as well as the nature of surface defects and residues.

Spectral Sensors: Spectral sensors record information on light outputs and fluctuations within the plasma area, which are reflective of the welding process's quality and consistency. This information enables one to fine-tune welding conditions to the most appropriate status.

These sensors are used together with a trigger sensor which helps in coordinating the welding and sensing systems in a way that all the data collected is real-time data. When these sensors are combined in a single system, they provide control and monitoring of the welding process, which in turn results in fewer weld defects.

6.2 CASE STUDY: NON-DESTRUCTIVE INSPECTION OF AIRCRAFT CORROSION

In this case study, a lap joint—an area susceptible to hidden corrosion—is inspected using a combination of:

Multifrequency Eddy Current (EC) and Pulsed Eddy Current (PEC): EC is responsive to surface and sub-surface discontinuities; the resolution of different frequencies is different. PEC determines the response of the material to the pulsed signals to get detailed information regarding the condition of the structure.

Digital Radiographic (X-Ray) System: Non-destructive testing is then followed by disassembling and cleaning of the lap joint for X-ray imaging. This destructive measurement is useful in ascertaining the accuracy of the fusion model as well as establishing the amount of corrosion.

Fusion Approaches:

Several data fusion methods are employed to analyze the corrosion:

Multiresolution Pixel-Level Fusion: This technique employs the previously defined basis functions to merge data at pixel level and thereby increases the signal-to-noise ratio and the quality of the results.

Classification-Based Approach: The measurements that are collected are divided into distinct bins depending on the ground truth from the X-ray image. This approach gives a number of material losses, though it gives discrete values rather than continuous values.

Generalized Additive Model (GAM): GAM is used to forecast future remaining thickness as continuous values using EC and PEC data. It gives a broader outlook on corrosion than the basic categorization.

Dempster-Shafer Theory: This method uses fuzzy k-mean clustering, Dempster-Shafer combination rules, and locally weighted regression to predict the remaining thickness.

It combines several measurement data to come up with continuous values with high precision.

The obtained fusion results are compared to the X-ray thickness map to show better accuracy with the RMSE being much lower than in other methods.

Relevance to Our Framework:

The following case studies and examples demonstrate the significance of sensor fusion in intelligent manufacturing, which is consistent with the goals of the presented framework. It is possible to obtain more precise and reliable manufacturing processes by integrating various sensor data and applying appropriate fusion methods. In line with such real-world implementations of the proposed framework, there are significant benefits in terms of process efficiency, quality, and interpretability across multiple industries based on our emphasis on multimodal sensor integration and explainable AI methods.

7 DISCUSSION

7.1 Interpretation of Results

The results from our research confirm the effectiveness of the proposed multimodal sensor fusion framework in enhancing intelligent manufacturing systems. By employing both early and attention-based fusion strategies, our approach significantly improves prediction accuracy. This is clearly demonstrated in our experiments with the Condition Monitoring of Hydraulic Systems dataset, where the Random Forest Classifier achieved perfect accuracy, with precision, recall, and an F1-score of 1.00 across all classes. Such results underscore the potential of our approach to deliver precise condition monitoring, which is crucial for predictive maintenance and fault detection in industrial systems.

However, these near-perfect classification results also raise a potential concern regarding overfitting—a common issue where a model performs exceptionally well on training data but struggles to generalize to new, unseen data. The high accuracy observed may indicate that the model is overly tailored to the specific characteristics of the training datasets. To address this, future work should focus on implementing cross-validation techniques across a broader range of datasets to better assess the model's robustness. Additionally, incorporating regularization methods or pruning strategies could help prevent overfitting by simplifying the model and reducing its sensitivity to noise in the training data. Ensuring that the proposed framework maintains its high accuracy across different manufacturing contexts without compromising generalizability is essential.

The integration of dynamic prioritization through attention mechanisms further enhances system reliability by focusing on the most relevant sensor inputs, ensuring optimal performance across various manufacturing conditions. This adaptability is crucial for maintaining robustness and efficiency, even in complex and changing environments.

Additionally, the evaluation of the **Gas Sensor Array Drift Dataset** highlighted the model's ability to manage sensor drift. By incorporating advanced feature engineering techniques and leveraging attention mechanisms, our framework dynamically adjusted to sensor changes, achieving RMSE and MAE values that demonstrate its robustness in handling drift-related challenges.

The hybrid fusion strategy, which combines early fusion with attention-based techniques, optimizes computational efficiency and improves predictive capabilities. The results, with a cross-validation score of approximately 0.996, demonstrate the robustness of the fusion strategy in accurately predicting system conditions.

These findings emphasize the benefits of integrating advanced machine learning techniques with sensor fusion strategies, offering a powerful tool for improving the reliability and accuracy of intelligent manufacturing systems. Future work should explore further validation across diverse datasets and investigate more sophisticated models to ensure the robustness and scalability of the proposed framework.

7.2 Industry Implementation and Benefits of the Proposed Framework

The application of the proposed framework employs state-of-the-art sensor fusion in order to achieve a system perspective of the manufacturing processes. The proposed framework uses different types of sensors such as vision, x-ray, laser, and spectral sensors to identify faults that are often physically hidden and hence hard to identify. For instance, in laser welding, the task is to monitor the process and detect shortcomings that might slow it down; these include metal spatters or incomplete fusion, and for this purpose, the facility uses UVV visual sensors, x-rays, and laser reflective photoelectric sensors. In addition to its use in the accurate identification of faults, this capability can be used to make corrections on the processes in manufacturing as soon as a fault is identified, thereby reducing time wastage on faults.

7.3 Advantages of the Proposed Framework

7.3.1 Enhanced Fault Detection and Process Optimization

The proposed framework uses advanced sensor fusion to provide a comprehensive view of manufacturing processes. The combination of different types of sensors, including visual and x-ray, laser, and spectral ones, provides real-time identification of hidden faults, which strengthens the fault identification process considerably. As an example, in laser welding technology use of sensors such as UVV visual sensors, x-ray, or laser reflective photoelectric sensors successfully flag problems such as metal spatters or inadequate fusion. This capability enables the management to make corrections promptly, hence improving the effectiveness of the process and reducing unnecessary shutdowns.

7.3.2 Improved Productivity and Reduced Cost

The application of the outlined framework improves the level of automation and addresses process control, therefore promoting the growth of product quality and lowering the percentage of defects. This results in huge cost-savings and assurance of correctness since the framework pinpoints defects as well as enhances manufacturing procedures. The manufacturing of inferior goods reduces both the time and energy that would have been incurred in the reprocessing of the same which reduces ‘time to market’.

7.3.3 Versatility and Broader Industrial Implementation

Most of the concepts outlined in the proposed framework can be easily adapted in the general manufacturing industries like automotive industries, electronics industries, etc., Its multimodal sensor fusion approaches and methods of AI explainability are extensible. For example, in the automotive sector, it can oversee the assembly line, and, in the electronics sector, it can enhance quality control and defect identification on printed circuit boards. Such flexibility also shows the applicability of this framework in various manufacturing issues across industries.

7.4 Limitations and Future Work

Despite the promising performance of the hybrid fusion model, several limitations need to be addressed:

Computational Demands: The use of multiple layers is a crucial feature of the hybrid fusion model, that increases the demands of computations. It becomes even more complicated, especially for applications that require real-time processing and, therefore, high processing capacity. Further studies to extend this work should be conducted to implement a more accurate and efficient framework by exploring the ways to reduce the computations, for example, through pruning, quantizing, or using better algorithms without loss of the performance of the fusion model.

Dataset Quality and Diversity: The training datasets are most critical to the performance and effectiveness of the framework in question. The performance of the recommended methodology was shown using the Condition Monitoring of Hydraulic Systems dataset for testing, but the framework presented can only be recommended with certain reservations regarding their applicability to other manufacturing situations. Future work should involve widening the pool of data so that it can contain data on more diverse types of manufacturing environments. Other efforts would be to look into procedures for applying transfer learning methods that, as much as possible, can fine-tune the model on knowledge of other similar tasks or fields.

Sensor Types and Data: The framework could be enhanced by using additional types of sensors, such as environmental and chemical sensors, to provide a more thorough view of the manufacturing environment. It would have a great possibility of providing better solutions and organization of work in the processes. Future works should also examine how other sensors could be incorporated into the system and whether they would help augment the result.

Sensor Drift: The problems associated with the sensor drift were illustrated on the Gas Sensor Array Drift dataset. Even when using intelligent feature engineering as well as attention-based fusion approaches, the model had certain challenges like drift over time. Future work should also integrate more refined drift detection techniques to adapt its settings change based on the sensor's variation, This could be achieved by using an ensemble version of the drift detection or integrating temporal models such as Recurrent Neural Networks (RNN) or Long Short-Term Memory (LSTM). Other possibilities include improving attention mechanisms that depend on the importance of the sensors and on temporal variations.

8 CONCLUSION

In conclusion, this work presents a new approach to embedding multimodal sensors into the IPS and the further utilization of computer vision and XAI for improving the functionality of the given systems. Our presented approach of having a deep learning model fused with attention mechanisms reported remarkable enhancements in effectiveness, error detection, flow enhancement, and quality assurance compared with the traditional methodologies.

The improvement of the proposed framework's accuracy is evident in our experiments' empirical results. On the Condition Monitoring of Hydraulic Systems dataset, the model demonstrated almost no impaired classification that would indicate low efficiency of fault detection. In the same way, when it comes to handling sensor drift in the Gas Sensor Array Drift dataset, it was learned that the model can be trained while showing some potential for improvement.

The framework tackles severe problems in data integration and analysis making the smart manufacturing systems fundamentally more sophisticated. To that end, it increases efficiency, reliability, and audibility: evidenced by its adoption in industrial contexts. Such results help to determine that this method can be applied to the present industries and has a potential for future developments of different types of industries. Therefore, the implementation of these AI technologies is expected to provoke major shifts in the advancement of manufacturing. Since it aims for the shift towards more efficient and more sustainable production methods, our framework can be regarded as a suitable instrument for pursuing the vision depicted by Kusiak (2019) and Monostori et al. (2016).

ACKNOWLEDGMENTS

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

FUNDING

Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not applicable.

ABOUT THE AUTHORS

ZHU, Yue

Georgia Institute of Technology, California, United States.

REFERENCES

[1] Liu, Z., Xiao, G., Liu, H., & Wei, H. (2024). Multi-Sensor Measurement and Data Fusion. IEEE Transactions on Instrumentation and Measurement, 73, 234-245. https://ieeexplore.ieee.org/abstract/document/9693406?casa_token=40VH55MYDtEAAAAA:5M5RuueqkCua6ONFCVjNeNl8V0TlZ4-giZ8VQzK0U6CXh_TefMxhQaCBotzrZXsBrmaDEjnq

[2] Diez-Olivan, A., Del Ser, J., Galar, D., & Sierra, B. (2019). Data Fusion and Machine Learning for Industrial Prognosis: Trends and Perspectives Towards Industry 4.0. Journal of Manufacturing Processes, 40, 123-134. https://www.sciencedirect.com/science/article/pii/S1566253518304706?casa_token=Bi4TM627bvIAAAAA:FIhyFRBhR4K3o7_Ly5thLLoHChk7UYVPCZTvMDEK5aPxTYUJQuAPOI_Dn3ZHRIRz1TUbIaQ_

[3] Helwig, Nikolai, Pignanelli, Eliseo, and Schtze, Andreas. (2018). Condition monitoring of hydraulic systems. UCI Machine Learning Repository. https://doi.org/10.24432/C5CW21.

[4] Vergara, Alexander. (2013). Gas Sensor Array Drift at Different Concentrations. UCI Machine Learning Repository. https://doi.org/10.24432/C5MK6M.

[5] Buscema, M., Terzi, S., & Tastle, W. (2010). Steel Plates Faults. UCI Machine Learning Repository. https://doi.org/10.24432/C5J88N

[6] Zhang, Z., Farnsworth, M., Song, B., Tiwari, D., & Tiwari, A. (2024). Deep Transfer Learning With Self-Attention for Industry Sensor Fusion Tasks. IEEE Transactions on Industrial Informatics. https://ieeexplore.ieee.org/abstract/document/9813440

[7] Sensor Data Fusion in Industry 4.0. (2024). Farnell. https://si.farnell.com/sensor-data-fusion-in-industry-40-trc-ar

[8] Goyal, D., & Pabla, B. S. (2015). Condition-based maintenance of machine tools—A review. CIRP Journal of Manufacturing Science and Technology, 10, 24-35. https://doi.org/10.1016/j.cirpj.2015.05.002

[9] Kang, H. S., Lee, J. Y., Choi, S., Kim, H., Park, J. H., Son, J. Y., ... & Do Noh, S. (2016). Smart manufacturing: Past research, present findings, and future directions. International Journal of Precision Engineering and Manufacturing-Green Technology, 3(1), 111-128. https://doi.org/10.1007/s40684-016-0015-5

[10] Kusiak, A. (2019). Fundamentals of smart manufacturing: A multi-thread perspective. Annual Reviews in Control, 47, 214-220. https://doi.org/10.1016/j.arcontrol.2019.02.001

[11] Liu, Y., Li, S., Liu, C., & Wang, H. (2021). Explainable AI for intelligent manufacturing: A survey. AI Open, 2(1), 105-123. https://doi.org/10.1002/aiopen.76

[12] Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., ... & Ueda, K. (2016). Cyber-physical systems in manufacturing. CIRP Annals, 65(2), 621-641. https://doi.org/10.1016/j.cirp.2016.06.005

[13] Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing, and service with big data. The International Journal of Advanced Manufacturing Technology, 94(9-12), 3563-3576.

https://doi.org/10.1007/s00170-017-0233-1

[14] Wuest, T., Weimer, D., Irgens, C., & Thoben, K. D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research, 4(1), 23-45. https://doi.org/10.1080/21693277.2016.1192517

[15] Yin, S., Ding, S. X., Xie, X., & Luo, H. (2014). A review on basic data-driven approaches for industrial process monitoring. IEEE Transactions on Industrial Electronics, 61(11), 6418-6428. https://doi.org/10.1109/TIE.2014.2301773

[16] Zheng, P., Wang, H., Sang, Z., Zhong, R. Y., Liu, Y., Liu, C., ... & Mubarok, K. (2018). Smart manufacturing systems for Industry 4.0: Conceptual framework, scenarios, and future perspectives. Frontiers of Mechanical Engineering, 13(2), 137-150. https://doi.org/10.1007/s11465-018-0499-5

[17] Zhong, R. Y., Xu, X., Klotz, E., & Newman, S. T. (2017). Intelligent manufacturing in the context of industry 4.0: a review. Engineering, 3(5), 616-630. https://doi.org/10.1016/J.ENG.2017.05.015

[18] Shao, Y., Yuan, X., Zhang, C., Song, Y., & Xu, Q. (2020). A Novel Fault Diagnosis Algorithm for Rolling Bearings Based on One-Dimensional Convolutional Neural Network and INPSO-SVM. Applied Sciences, 10(12), 4303. https://doi.org/10.3390/app10124303

[19] Qiu, S., Cui, X., Ping, Z., Shan, N., Li, Z., Bao, X., & Xu, X. (2023). Deep Learning Techniques in Intelligent Fault Diagnosis and Prognosis for Industrial Systems: A Review. Sensors, 23(3), 1305. https://doi.org/10.3390/s23031305