[1]Aashish Arora

[2]Rajeev Gupta

Intelligent and Automated Product Quality Assurance System using Machine Learning

Journal of Electrical Systems

Abstract: - The paper discusses the use of advanced technologies like the Internet of Things (IoT) and cyber-physical systems (CPS) to enhance product quality assurance in production industries. It proposes an intelligent approach using machine learning and termed the model as ETLViT, which integrates transfer learning models and vision transformers (ViT). The proposed model is a three-layered architecture integrated with a cyber-physical system (CPS). The first layer collects images of the processed products using a CPS, while the second layer transmits the data to a cloud server. The third layer analyzes the input image and predicts any faults in the product, with results also stored in the cloud server. The learning process involves noise removal, feature extraction, and classification, which are performed simultaneously by the ensemble model called ETLViT, designed using transfer learning and vision transformers. The result was presented with transfer learning models, ensemble model with machine learning and ETLViT. Among all of them, ETLViT outperforms the best.

Keywords: Automation, Product Quality Assurance, Fault detection, Machine Learning, Vision Transformer.

1. Introduction

The technique of examining arriving data in order to identify and address issues in advance is known as predictive quality assessment. By effectively detecting the core reason behind losses as well as averting losses in the future, this newly developing area of AI allows producers or manufacturers to reduce losses with respect to quality and waste. By using technologies from Industry 4.0, including predictive quality, production organisations can overcome a variety of quality assurance issues[1-4]. Computerised root cause assessment, predictive insights, as well as notifications and alerts in real-time are the three primary predictive quality capabilities (Fig. 1). Computerised root-cause evaluation and prediction assurance technologies can be used by manufacturers to find historically undiscovered root causes and continually analyse massive volumes of information in order to offer precise assessment of any issue. Additionally, data-driven insights from predictive quality can help manufacturers reduce issues with quality. Predictive maintenance, for example uses efficiency monitoring techniques to identify potential flaws and fix them prior to they occur. Contrarily, preventative maintenance emphasises regular maintenance procedures to lessen the likelihood of downtime and equipment failure [3][6]. Predictive quality can also provide real-time warnings and announcements that let the production staff know right away as any manufacturing inadequacies are found. This enables them to address the problem before it has an impact on output. Information related to manufacturing activities has grown at a never-before-seen rate as a result of Industry 4.0 along with Internet of Things developments in digitalization. As a result, there are now more opportunities for data-driven techniques like Machine Learning (ML) for tracking the processes of manufacturing [1]. Machine learning (ML) technologies that effectively employ the advantages of ML in mass production have already developed. The requirement of knowledge from production & data science is a basic barrier to scaling ML-based projects [2]. The implementation of machine learning (ML), that has been beneficial for a variety of purposes in manufacturing sector [1, 2], also a popular strategy for maximizing the potential of datasets. Direct collaboration for the advancement of technology is being done via Industry 4.0. Every day, decision-making requiring a vast input of data and personalization in the process of production confronts both robots and supervisors. One of the key difficulties in this field is the ability to foresee when

[1] [1]Research Scholar, M.M. Institute of Computer Technology & Business Management, Maharishi Markandeshwar (Deemed to be University) Mullana, Ambala, Haryana, India .

[2]Department of Computer Science and Engineering M.M. Engineering College, Maharishi Markandeshwar (Deemed to be University) Mullana, Ambala, Haryana, India .

Corresponding author mail : aashish.aroura@outlook.com

Second author mail : rajeev.gupta@mmumullana.org

upkeep for assets would be necessary. The ability to undertake predictive maintenance improves machine uptime, expenditures, management, and production grade [3].

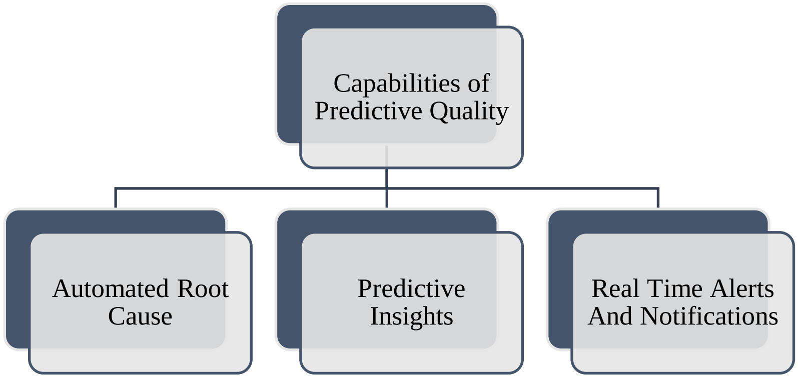

Figure 1 Capabilities of Predictive Quality

In the industrial industry, there has been rising curiosity in the combined use of machine learning methods with data visualization in automated maintenance. Numerous studies have recently looked into the potential benefits of these innovations to enhance maintenance processes, decrease downtime, and boost efficiency. Two examples of this type of research are Cheng et al.'s (2022) [5] comprehensive review of the literature on visual analytics as a form of predictive maintenance in the identical industry as well as Nacchia et al.'s (2021) [4] systematic representation of the developing deployment of ML methods to conduct predictive maintenance in the industry of manufacturing. various studies seek to provide a thorough overview of the state-of-the-art in various subjects today, to spot trends and problems, and to point up promising areas for further study. The state-of-the-art in this field is reviewed by Md et al. (2022) [6], who also emphasize how algorithms on machine learning may be used to raise the standard of manufacturing in the overall framework of Industry 4.0. They examine different data-driven approaches and techniques that are able to be used to forecast and manage quality metrics and they highlight the main obstacles and openings for further study. In the context of Industry 4.0, Lee et al. (2019) [7] suggest a quality control ecosystem for automated maintenance. To create a more comprehensive and proactive strategy for quality management, researchers underline the significance of integrating a number of technologies as well as information sources that involve sensors, Internet of Things (IoT) devices, along with data analysis. They also emphasize how predictive management has the ability to increase machinery effectiveness and decrease downtime. Employing Matlab tools, Hrehova (2016) [8] suggests a prediction model to assess the manufacturing method's efficiency. The model may be used to track and examine many quality metrics in real-time because it relies on statistical process control (SPC) methodologies. Therefore, the key contribution of the paper are:

- In this paper, three-tier architecture is proposed including cyber-physical systems and cloud for product quality assurance.
- The methodology explore the ensemble approach of vision transformer with machine learning for defect prediction in manufacturing products.
- The paper also presented the comparative result evaluation with some state-of-art models as well as with existing automated quality assurance system.

The rest of the paper is organized as: Section 2 presents the literature review for the analyzing the models presented for product quality assurance. Section 3 presents the methodology used in this paper. This section gives detailed description about the model adopted and algorithms. Section 4 presents the result analysis with comparative state-of-art. Section 5 presents the current limitations of work and future research directions. Finally in section 6 conclusion of the paper is presented.

2. Literature Review

Traditionally, manufacturers have relied on quality management systems to ensure that their products meet the essential standards. This often involves inspecting a sample of the products to verify for any faults or non-conformities at the end of the production operation. However, there are a couple problems with this tactic. To begin with, it is ineffective at avoiding issues from happening altogether and is only good for detecting them once they have taken place. Particularly for large-scale businesses, checking every manufactured product manually can be laborious, costly, and susceptible to human mistake. The implementation of data and analytics to identify developments and patterns in the manufacturing procedure, on the other hand, enables producers to foresee and thwart quality issues prior to the materialize. The application of ML techniques for predicting quality in manufacture is the topic of the studies by Nalbach et al. (2018) [9] as well as Saadallah et al. (2022) [10]. A novel theory of quality assurance that makes use of ML techniques for predicting quality is put forth by Nalbach et al. (2018) [9]. They provide an architecture that includes data collection from diverse sources, cleaning of data and preparation, machine learning-trained predictive models, and lastly, the deployment of these models to anticipate future quality concerns in real-time. On the other hand, Saadallah et al. (2022) [10] concentrate on the creation of a system of understandable predictive quality checks for the manufacture of electronic products. They first extract information from photos of the electronic parts using deep learning methods, and then they utilize a decision tree algorithm to categorize the elements depending on whether they are defective or not. By highlighting the crucial elements that influenced the model's choice, a gradient-based approach is used to increase the comprehensibility of the model. In order to accomplish this, it is possible to analyze data from a variety of sources such as sensor readings, machine logs, and quality control audits. By using AI and ML algorithms for recognizing them, manufacturers may spot trends and anomalies in data that indicate potential quality issues and take pre-emptive action.

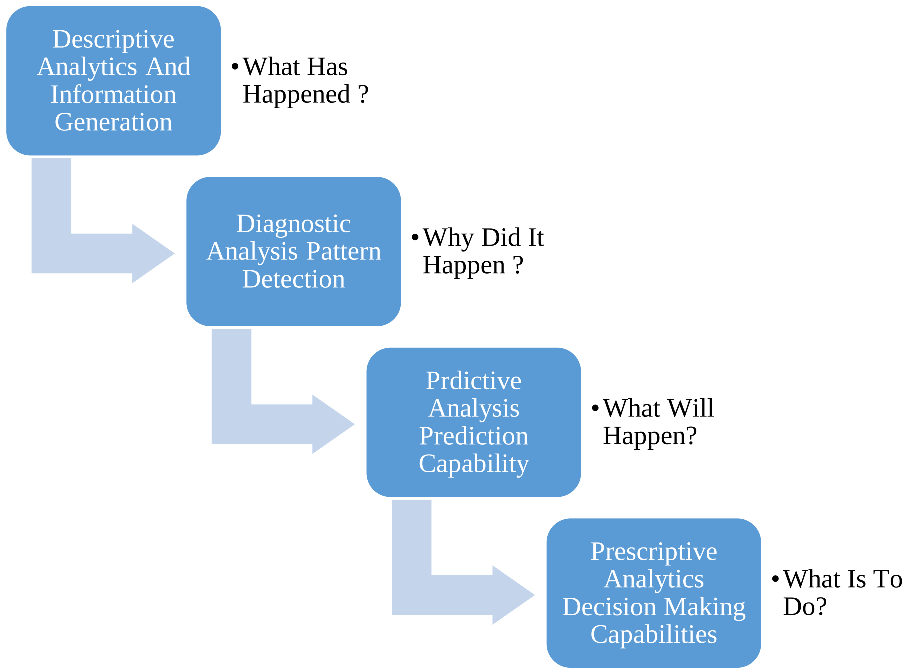

Figure 2 Predictive Quality Analysis

Diagnostic evaluation entails determining and comprehending the root cause of an already-occurring event. This is achievable using pattern-detection algorithms that assist in identifying the fundamental causes of the occurrence. Employing data and statistical computations, predictive analysis determines the possibility that a future event will occur. This aids in identifying the incident's primary cause and enables preventative measures to be taken to stop it from occurring happening in the future. The predictive analysis's capacity to forecast enables us to base our decisions on the knowledge we have obtained from studying past data. Recognizing the events that are likely to occur in the future allows us to be proactive in preventing potential issues or seizing new possibilities [11–15]. A predictive analytics strategy for ensuring the quality of steel billets in a manufacturing process was put forth by Belov et al. in 2022 [16]. A hybrid model for product-process-machine needs was created by Voisin et al. (2018) [17] for machine tool maintenance planning as well as part quality control. A data-driven strategy for enhancing multistage production processes utilizing predictive maintenance and quality assurance analysis was described by Cui et al. in 2022[18]. The difficulties and prospects of Quality 4.0 were examined by Zonnenshain and Kenett (2020) [19] in terms of quality engineering. Production statistics and industrial IoT play a part in enhancing the calibres of production processes, according to Lade et al. (2017) [20]. The opportunities and difficulties of quality engineering for additive manufacturing methods have been identified by Colosimo et al. (2018) [21]. The enabling methods, concepts, and potential uses of predictive manufacturing were given by Pulikottil et al. (2021) [22]. The essential role of Quality 4.0 in improving the industrial sector completely was highlighted by Javaid et al. (2021) [23]. In order to attain environmental sustainability Kumar et al. (2021) [24] suggested predictive data analysis for managing energy in an intelligent manufacturing facility. In the manufacturing sector, Alok et al. (2018) [25] investigated the empirical prediction of behavioural intention toward implementing lean approaches.

3. Methodology

In this paper, ensemble model is prepared using transfer learning models with vision transformer (ViT) for detecting abnormalities in products. This section describes the architecture of the designed model. The article highlights the challenges faced by the industry in detecting tiny defects like missing screws during manual processing and the overwhelming quantity of products that need to be processed for quality control. To address these issues, the authors propose a system that utilizes images of the end product taken at the production line to detect defects. The model presented in the paper is a three-layered architecture integrated with a cyber-physical system. The first layer is responsible for collecting images of the processed products using a cyber-physical system (CPS). The second layer transmits the collected data to a cloud server. Finally, the third layer is dedicated to analysing the input image and predicting any faults in the product. The results are also stored in the cloud server. This learning process is divided into three steps: noise removal, feature extraction, and classification. To accomplish this, an ensemble model is designed with transfer learning and vision transformer (ETLViT) has been designed. This model is capable of performing these processes simultaneously. Figure 3 illustrates the architecture of the ensemble model ETLViT.

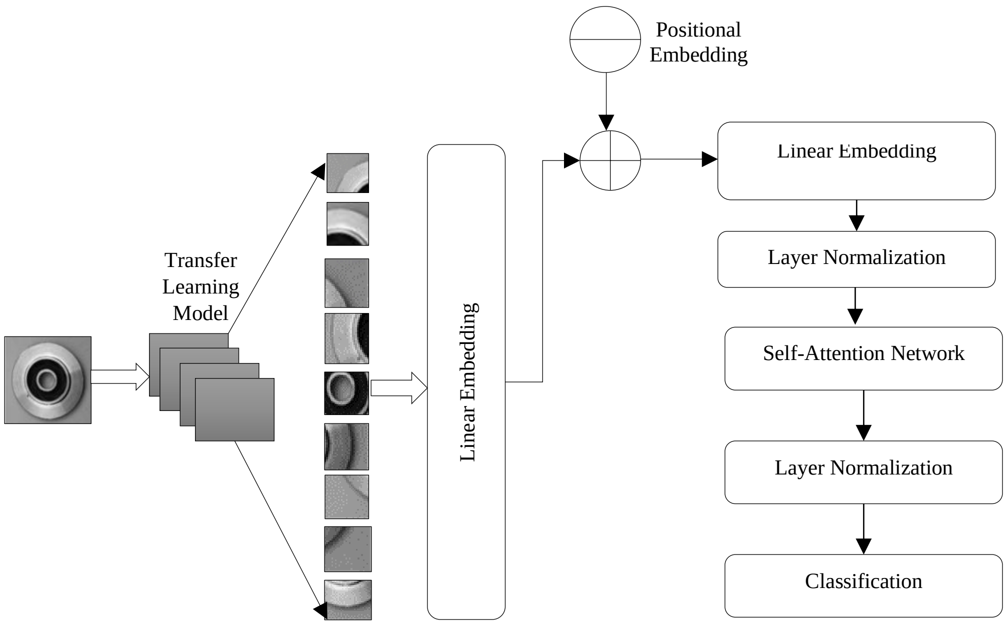

Figure 3 Proposed Ensemble Model for Quality Assurance

The input images in the study were resized to dimensions of 128x128, and pre-trained weights from the ImageNet dataset were utilized. A batch size of 64 was consistently employed during the training process. The ETLViT model converts the images into 1D-array feature maps. For multiclass classification, the final layer of the model employs the softmax activation function. The Adam optimizer with a learning rate of 1e-5 was utilized. The Adam optimizer dynamically adjusts the learning rates for individual parameters using an adaptive learning rate technique:

$$l = -\sum_{c=1}^{M} y_{o,c} \log (p_{o,c}) \tag{1}$$

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \tag{2}$$

Where σ is termed as softmax function. Here input vector is represented as z, and the exponential function as $z(i)$ represents the standard exponential function for the ith feature vector. The output vector consists of C classes, where the exponential function of $z(j)$ represents the standard exponential function. For transfer learning, VGG19, Resnet50, and Densenet201 models were selected for ensembling with ViT. The ensemble model employed in this study, which combines transfer learning and vision transformer, offers several advantages. Firstly, it improves performance by leveraging the strengths of multiple models. Secondly, it enhances the robustness and generalization of the model by reducing errors and capturing diverse patterns. Additionally, the ensemble model utilizes complementary information obtained through different transformations, making it more effective in handling vision-based tasks. It also takes advantage of pre-trained models for faster initialization, particularly useful when data is scarce. Finally, the ensemble model demonstrates enhanced resistance to perturbations through the use of data augmentation techniques. These are discussed below:

3.1 VGG19

VGG19 is a popular transfer learning model used in this study. It is a convolutional neural network model with a total of 19 layers. Out of these layers, 16 are convolutional layers, 3 are fully-connected layers, and there are 5 max-pooling layers along with a softmax layer. The convolutional layers have a kernel size of 3x3 and a stride value of 1. Spatial padding is also applied to preserve the resolution of the image. The ReLU activation function is used after each convolutional layer in VGG19, which helps maintain the non-linearity of the model. By default, the input image size for VGG19 is set to 224x224 pixels with 3 RGB channels. The spatial dimensions of the processed features are reduced through the use of max-pooling layers with a kernel size of 2x2 and a stride value of 2. The last fully-connected layer in VGG19 consists of 4096 neurons, and it is connected to the softmax layer for classification. The VGG19 model has a total of 19.6 billion learning parameters.

3.2 ResNet50

ResNet is a deep neural network that has achieved high performance in various image recognition competitions. It addresses the limitations of traditional CNNs by introducing shortcut connections, allowing direct mapping from input to output. This architecture, consisting of residual blocks, is widely used for image classification, object localization, and detection tasks. ResNet models have simpler structures compared to VGG networks and fewer filters. The ResNet architecture includes skip connections that connect the input directly to the output, enabling the learning of residual mappings. This overcomes the issue of vanishing gradients and improves the performance of deeper layers. The ResNet50 model, in particular, uses a bottleneck architecture and consists of several convolutional and pooling layers. It starts with a 7x7 convolutional layer with 64 kernels and a stride size of 2, followed by a max-pooling layer with a stride of 2. There are three sets of convolutional layers with varying kernel sizes and repeat factors. After the convolutional layers, an average pooling layer is applied, followed by a fully connected layer with 1000 nodes and a softmax function.

3.3 Densenet201

DenseNet is a fully connected convolutional neural network (CNN) architecture that introduces dense connections between layers. Unlike traditional CNNs, DenseNet connects each layer to all preceding and subsequent layers, forming dense connections. This design addresses the vanishing gradient problem and improves accuracy in deep neural networks. DenseNet shares similarities with ResNet but incorporates distinctive features. It is known for its parameter efficiency, utilizing a small number of parameters per layer. Additionally, DenseNet benefits from deep implicit supervision, which improves the flow of gradients throughout the network. The architecture of DenseNet consists of dense connections, that exchange information across multiple CNN layers. Unlike traditional CNNs, DenseNet concatenates the feature maps that will reduce the number of training parameters. The base structure of DenseNet composed of layers such as batch normalization, ReLU activation, and 3x3 convolutions at each step. It consists of dense blocks with varying number of layers. Each layer within a dense block consists of a bottleneck layer with a 1x1 kernel size, followed by a convolutional layer with a 3x3 kernel size. Additionally, an average pooling layer with a stride of 2 is included.

3.4 Vision Transformer

Let $S = \{X_i, y_i\}_{i=1}^{r}$ donate a set of r images, where X_i is an image and y_i is its corresponding class label $y_i \in \{1, 2, \ldots, \mathrm{m}\}$ and m is the number of defined classes for that set. The objective of the Vision Transformer model is to learn the mapping from the sequence of image patches to the corresponding semantic label. The Vision Transformer is an architecture inspired by the Transformer model, which has gained popularity for its excellent performance in natural language processing tasks. The Transformer uses an encoder-decoder architecture and can process sequential data efficiently without the need for recurrent networks. One key feature of the Transformer is the self-attention mechanism, which allows it to capture long-range dependencies between elements in a sequence. The Vision Transformer adapts this architecture for image-related tasks, leveraging self-attention to capture relationships between image patches and achieve state-of-the-art performance in computer vision. The Vision Transformer is an extension of the Transformer model for image classification tasks. Its main objective is to apply the Transformer architecture to non-textual data without incorporating any domain-specific design. The Vision Transformer employs the encoder module of the Transformer to perform classification by mapping a sequence of image patches to semantic labels. Unlike traditional convolutional neural network (CNN) architectures that use filters with a local receptive field, the Vision Transformer utilizes the attention mechanism to attend to various regions of the image and integrate information across the entire image. The Vision Transformer consists of an embedding layer, an encoder, and a final head classifier. In the first step, an image is divided into non-overlapping patches. Each patch is treated as an individual token by the Transformer. For an image of size $c \times h \times w$ (where h is height, w is width, and c is the number of channels), patches of size $c \times p \times p$ are extracted. This creates a sequence of patches $(x1, x2, \ldots, xn)$ with a length of n, where $n = hw/p\text{^}2$. The patch size p is typically chosen as $16 \times 16\ or\ 32 \times 32$, with smaller patches resulting in longer sequences and vice versa.

Linear Embedding Layer : To prepare the input for the encoder, the sequence of image patches undergoes linear projection using a learned embedding matrix E, resulting in embedded representations. These embeddings are concatenated along with a learnable classification token v_{class}, which is necessary for the classification task. The Transformer treats the embedded image patches as a set without considering their order. To preserve the spatial arrangement of the patches from the original image, positional information (encoded as E_{pos}) is added to the patch representations. The final input to the encoder is the concatenated embedded sequence of patches along with the token z_0, which can be represented as:

$$z_0 = [v_{class}; x_1E; x_2E; ,,,,,,,,,,, x_nE;] + E_{pos}, E \in R^{(p^2c)x\,d}, E \in R^{(n+1)x\,d} \tag{3}$$

Either 1-D or 2-D positional encodings yields similar outcomes. As a result, a straightforward 1-D positional encoding method is employed to maintain the positional information of the flattened image patches.

Vision Transformer Encoder : The embedded patches z_0 are fed into the Transformer encoder. The encoder consists of L identical layers consisting of self-attention block and a fully connected feed-forward dense block represented as:

$$z_l' = MSA\big(LN(z_l - 1)\big) + z_{l-1} \qquad l = 1 \ldots\ldots\ldots\ldots\ldots\ldots.l \tag{4}$$

$$z_l = MLP\left(LN(z_l')\right) + z_l', \qquad l = 1 \ldots\ldots\ldots\ldots\ldots\ldots . l \tag{5}$$

At the last layer of the encoder, we take the first element in the sequence z_0^t and pass it to an external head classifier for predicting the class label.

$$Y = LN(z_0^t) \tag{6}$$

Training algorithm is presented below as algorithm 1. For experiment, we have used python for training and testing purpose. For this the entire dataset is divided into 70:30 ratio and trained. First of all the denoising CNN is trained and further denoised image feature samples are fed into ResNet50 for further training. For training, the input image size is taken as $\mathbf{64 * 64}$. The minimum batch size for training was taken to be 64. The maximum epoch was taken to be 100. For training denoising CNN, mean square error (MSE) was taken as loss function. Notably, these model utilizes approx. 10GB RAM on a high-performance GPU computing. Therefore, this model was trained on computing service provided by google i.e., google colab.

Algorithm 1 Training Process

Input: I_i, Input images;
Training dataset, $T_d = \{I_i^n\}, where\ n \in size(T_d)$

Output: O_i^n, Product Type

1: Initialization: $epoch_{maxdenoise}$ {maximum epoch for denoising), $epoch_{maxcnn}$ {maximum epoch for detection}
2: For i=1: $epoch_{maxdenoise}$
3: DI_i←*DenoisingCNN*(I_i), *DI_i is the denoised image*
4: For j=1:$epoch_{maxcnn}$
5: O_i^n← ETLViT DI_i), *DI_i is the denoised images*
6: While $loss$ reaches convergence do
7: Minimize($loss$)
8: end
9: Return O_i^n

4. Results and Discussion

This section focuses on the implementation details, result analysis, and a comparison with state-of-the-art models for quality assurance. The designed model is implemented on the Tesla P100-PCIE GPU provided by Google Colab, using Keras and Tensorflow as the backend. Sub-section 4.1 provides a description of the dataset used in the implementation. In Section 4.2, performance evaluation metrics are discussed, explaining how the model's performance was measured. Subsection 4.3 presents the result analysis, providing detailed information about the performance of the designed model. This section also presented the comparative analysis, comparing the proposed model with existing state-of-the-art models for quality assurance. This comparison helps to assess the effectiveness and superiority of the designed model in detecting abnormalities in products, highlighting its advantages over other approaches.

4.1 Dataset Description

This paper focuses on using data samples of defective and non-defective casting product images for quality inspection. The dataset is sourced from Kaggle [26][27] and specifically involves images of submersible pump impellers. Submersible pump impellers are vital components that enable fluid circulation in pumps. When fluids enter the centrifugal pump, they flow into the center of the impeller, and as the impeller rotates, the velocity of the fluid increases, ultimately pushing it outwards through centrifugal force. The impeller's role is crucial in pump functionality, highlighting the criticality of high-quality casting for submersible pump impellers. The study presents images of these impellers, including samples of defective products and normal product images. Additionally, the paper employs another dataset, the NEU Metal Surface Defects Database [28], for performance analysis. This dataset contains 1,800 grayscale images showcasing six typical surface defects found on hot-rolled steel strips, namely rolled-in scale, patches, crazing, pitted surface, inclusion, and scratches

4.2 Performance Evaluation Parameters

The designed model is evaluated on following metrices:

$$\text{Accuracy} = (\text{TrP} + \text{TrN})/(\text{TrP} + \text{TrN} + \text{FlP} + \text{FlN}) \tag{2}$$

$$\text{Recall / Detection Rate} = \text{TrP} / (\text{TrP} + \text{FlN}) \tag{3}$$

$$\text{Precision} = \text{TrP} / (\text{TrP} + \text{FlP}) \tag{4}$$

$$\text{F1\_Score} = 2^* (\text{Precision}^* \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{5}$$

Where, TrP = True Positive (condition when both actual and predicted values are of defective quality).

TrN = True Negative (Condition when both actual and predicted values are of good quality).

FlP = False Positive (Condition when actual is defective and predicted is of good quality).

FlN = False Negative (Condition when actual is of good quality and predicted is defective).

4.3 Result Analysis

The table 1 provides a performance comparison of different transfer learning (VGG19, ResNet50, and Densenet121), ensemble model of transfer learning and machine learning (ResNet50+SVM, and ResNet50+k-NN) and ETLViT (Vgg19+ViT, Resnet50+ViT, and Densenet121+ViT) for Casting Product Quality Inspection. The metrics evaluated include accuracy, precision, recall, and F1-score. Based on the table, it can be observed that the Vgg19 and Densenet121 models achieve an accuracy of 0.84, while the Resnet50 model achieves a slightly higher accuracy of 0.89. The ensemble models, such as Resnet50+SVM and Resnet50+k-NN, perform even better with accuracies of 0.93 and 0.92, respectively. Additionally, the models combined with Vision Transformers (ViT), such as Vgg19+ViT, Resnet50+ViT, and Densenet121+ViT, demonstrate exceptional performance with accuracies, precision, recall, and F1-scores of 0.99, indicating highly accurate predictions across the board. In summary, the models Vgg19+ViT, Resnet50+ViT, and Densenet121+ViT exhibit the highest performance across all evaluation metrics. These models achieve a remarkable accuracy, precision, recall, and F1-score of 0.99, indicating exceptional predictive capabilities. The table 2 provides a performance comparison of different transfer learning (VGG19, ResNet50, and Densenet121), ensemble model of transfer learning and machine learning (ResNet50+SVM, and ResNet50+k-NN) and ETLViT (Vgg19+ViT, Resnet50+ViT, and Densenet121+ViT) for metal surface quality inspection. Based on the results, the models Vgg19+ViT, Resnet50+ViT, and Densenet121+ViT achieve the highest accuracy and overall performance with scores of 0.98 across all metrics. The combination of Vision Transformers with these models enhances their performance, resulting in highly accurate and reliable predictions. If accuracy and overall performance are the primary considerations, these models should be preferred for machine learning tasks. Below in figure 4(a) represents the confusion matrix of ETLViT models designed using Vgg19+ViT, Resnet50+ViT, and Densenet121+ViT. Figure 4(b) represents the training and validation accuracy of the model and it was observed that all three models with vision transformer improvised the result. Similarly, fig 4(c) represents the training and validation loss of the model and it was observed that all three models with vision transformer improvised the result.

Table 1. Performance Comparison for Casting Product Quality Inspection

Models	Accuracy	Precision	Recall	F1-score
Vgg19	0.84	0.85	0.85	0.85
Resnet50	0.89	0.88	0.89	0.89
Densenet121	0.84	0.84	0.83	0.84
Resnet50+SVM	0.93	0.94	0.93	0.93
Resnet50+k-NN	0.92	0.92	0.91	0.91
Vgg19+ViT	0.99	0.99	0.99	0.99
Resnet50+ViT	0.99	0.99	0.99	0.99
Densenet121+ViT	0.99	0.99	0.99	0.99

Table 2. Performance Comparison on Metal Surface Quality Inspection

Models	Accuracy	Precision	Recall	F1-score
Vgg19	0.82	0.81	0.76	0.78
Resnet50	0.79	0.88	0.81	0.84
Densenet121	0.85	0.85	0.86	0.85
Resnet50+SVM	0.96	0.96	0.96	0.96
Resnet50+k-NN	0.96	0.96	0.96	0.96
Vgg19+ViT	0.98	0.94	0.95	0.95
Resnet50+ViT	0.98	0.97	0.98	0.98
Densenet121+ViT	0.98	0.98	0.98	0.98

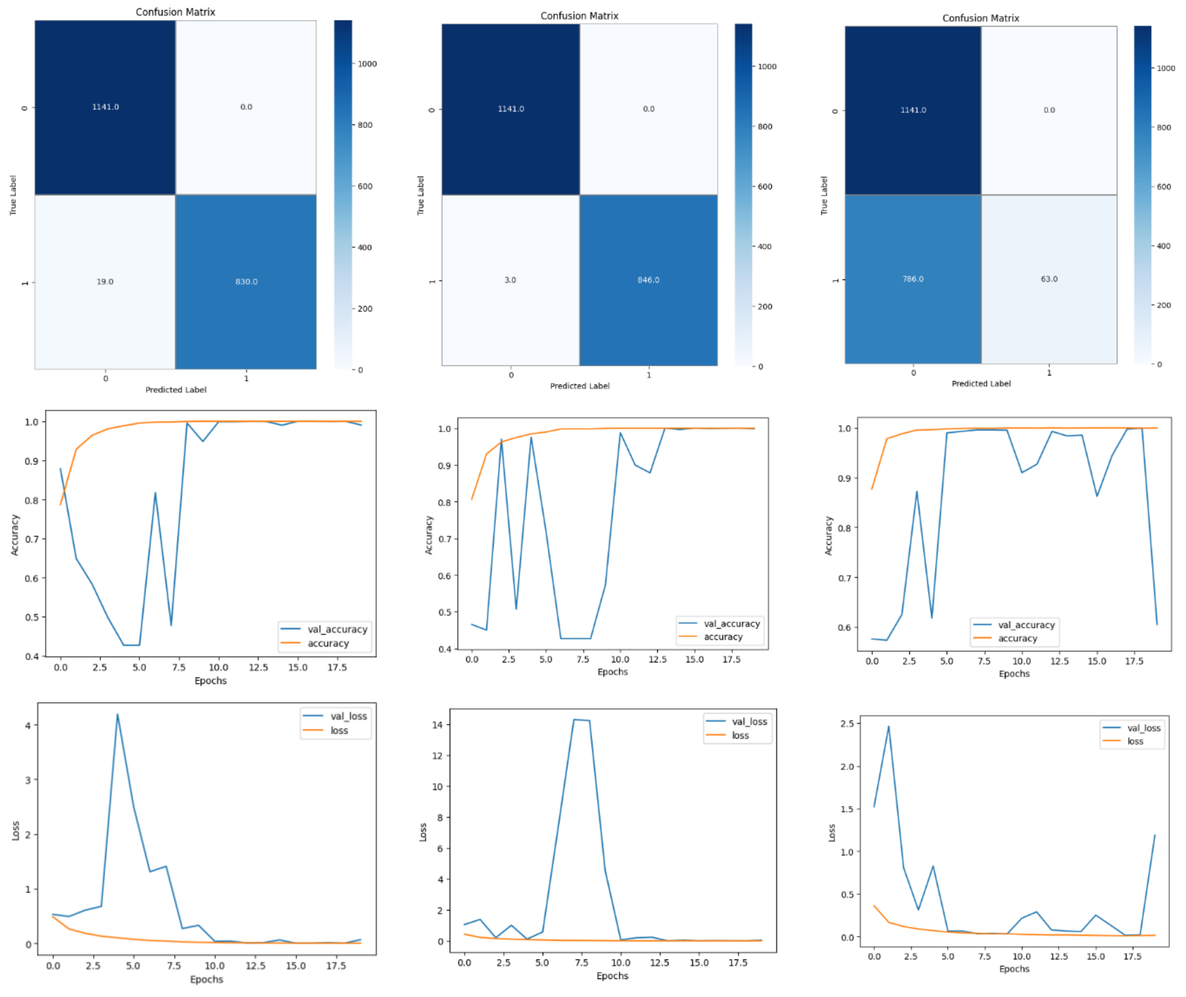

Figure 4 Training and Validation Performance of ETLViT for Quality Assurance

Figure 5 shows a comparative analysis of automated quality assurance with existing state-of-art techniques. For the gradient, Boosted Tree [29] accuracy is 93%, For CNN [30] accuracy is 86%, for XGBoost [31] accuracy is 93%, for CNN [32] accuracy is 94%, for CNN [33] accuracy is 95% and the proposed model (ETLViT) has maximum accuracy of 99%. The table shows that the ETLViT model has the highest accuracy score of 99%. This suggests that the ETLViT model outperforms the other state-of-the-art models in automated quality assurance and it can be inferred that ETLViT ensures quality assurance model has achieved more robust results.

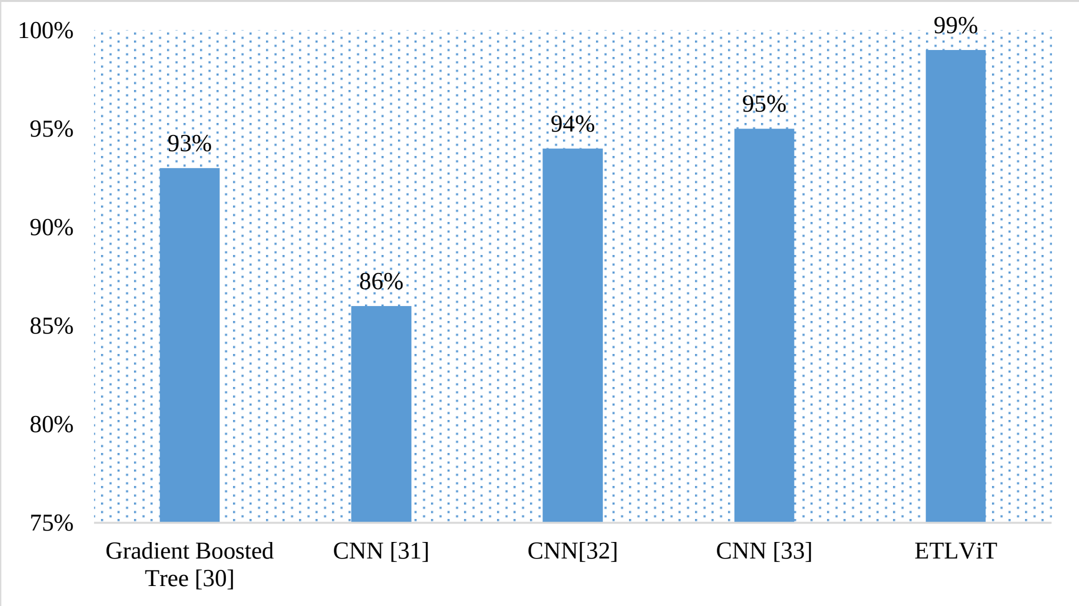

Figure 5. Comparative Analysis of Automated Quality Assurance State-of-Art Models

Table 3. Feature Comparison

Models	Intelligent	Automated	Cloud Storage	Noise Removal
[30]	GB	Yes	No	No
[31]	CNN	Yes	No	No
[32]	CNN	Yes	No	No
[33]	CNN	Yes	No	No
Proposed (ETLViT)	ETLViT	Yes	Yes	Yes

The feature comparison in Table 3 effectively justifies the proposed work presented in the paper through the distinct capabilities of the proposed ETLViT model. Unlike other models [30]-[33] which primarily utilize Gradient Boosting or Convolutional Neural Networks and lack automation and advanced features, the ETLViT model stands out. The proposed methodology is marked as “Intelligent”, indicating its use of machine learning techniques, likely combining transfer learning and vision transformers for enhanced decision-making and problem-solving. Moreover, the ETLViT model is designated as “Automated”, highlighting its ability to operate with minimal human intervention, a crucial aspect for efficient and consistent quality assurance. The inclusion of cloud storage and noise removal features further emphasizes its suitability for modern production environments, where handling large data sets and ensuring accuracy in quality assessment are dominant. Therefore, in summary, the ETLViT model’s advanced features shows itself as superior, intelligent, and automated solution for product quality assurance using machine learning.

5. Conclusion

In conclusion, this paper highlights the importance of implementing cyber-physical systems (CPS) in production industries to enhance and make it intelligent automated quality assurance processes. By integrating CPS into quality assurance procedures, businesses can achieve consistent production processes. Moreover, the use of CPS enables significant time savings by automating various aspects of quality control that would otherwise require manual labor and inspection. To combat this, the paper has presented a machine learning based approach to enhance quality assurance in production industries by leveraging cyber-physical systems (CPS) and advanced technologies such as transfer learning models and vision transformers. The proposed methodology, known as ETLViT, is a three-layered architecture that integrates CPS with an ensemble model for defect detection in processed products. The first layer of the architecture collects images of the processed products through the use of CPS, ensuring real-time data acquisition. The second layer facilitates the transmission of collected data to a cloud server, enabling seamless access and storage. Finally, the third layer performs image analysis and fault prediction, utilizing the ensemble model to simultaneously execute noise removal, feature extraction, and classification tasks. The paper presented the performance comparison of various transfer learning (VGG19, ResNet50, and Densenet121), ensemble model of transfer learning and machine learning (ResNet50+SVM, and ResNet50+k-NN) and ETLViT was conducted. The result showed that the combination of Vision Transformers (ViT) with various models, including Vgg19+ViT, Resnet50+ViT, and Densenet121+ViT, demonstrated better performance as compared to others. Therefore, the integration of Vision Transformers further enhanced the performance of these models, resulting in accurate and reliable predictions. Future research directions may focus on further optimizing the ensemble model, exploring different configurations of the CPS integration, and investigating it in real industrial environments.

References

[1] Zhou, B., Pychynski, T., Reischl, M., Kharlamov, E., & Mikut, R. (2022). Machine learning with domain knowledge for predictive quality monitoring in resistance spot welding. *Journal of Intelligent Manufacturing*, *33*(4), 1139-1163.

[2] Krauß, J., Pacheco, B. M., Zang, H. M., & Schmitt, R. H. (2020). Automated machine learning for predictive quality in production. *Procedia CIRP*, *93*, 443-448.

[3] Zonta, T., Da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the Industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, *150*, 106889.

[4] Nacchia, M., Fruggiero, F., Lambiase, A., & Bruton, K. (2021). A systematic mapping of the advancing use of machine learning techniques for predictive maintenance in the manufacturing sector. Applied Sciences, 11(6), 2546.

[5] Cheng, X., Chaw, J. K., Goh, K. M., Ting, T. T., Sahrani, S., Ahmad, M. N., ... & Ang, M. C. (2022). Systematic Literature Review on Visual Analytics of Predictive Maintenance in the Manufacturing Industry. Sensors, 22(17), 6321.

[6] Md, A. Q., Jha, K., Haneef, S., Sivaraman, A. K., & Tee, K. F. (2022). A Review on Data-Driven Quality Prediction in the Production Process with Machine Learning for Industry 4.0. Processes, 10(10), 1966.

[7] Lee, S. M., Lee, D., & Kim, Y. S. (2019). The quality management ecosystem for predictive maintenance in the Industry 4.0 era. International Journal of Quality Innovation, 5, 1-11.

[8] Hrehova, S. (2016). Predictive model to evaluation quality of the manufacturing process using Matlab tools. Procedia Engineering, 149, 149-154.

[9] Nalbach, O., Linn, C., Derouet, M., & Werth, D. (2018). Predictive quality: towards a new understanding of quality assurance using machine learning tools. In *Business Information Systems: 21st International Conference, BIS 2018, Berlin, Germany, July 18-20, 2018, Proceedings 21* (pp. 30-42). Springer International Publishing.

[10] Saadallah, A., Büscher, J., Abdulaaty, O., Panusch, T., Deuse, J., & Morik, K. (2022). Explainable Predictive Quality Inspection using Deep Learning in Electronics Manufacturing. *Procedia CIRP*, *107*, 594-599.

[11] Burggräf, P., Wagner, J., Heinbach, B., Steinberg, F., Schmallenbach, L., Garcke, J., ... & Wolter, M. (2021). Predictive analytics in quality assurance for assembly processes: lessons learned from a case study at an industry 4.0 demonstration cell. *Procedia CIRP*, *104*, 641-646.

[12] Saihi, A., Awad, M., & Ben-Daya, M. (2023). Quality 4.0: leveraging Industry 4.0 technologies to improve quality management practices–a systematic review. *International Journal of Quality & Reliability Management*, *40*(2), 628-650.

[13] Wei, Y., Wu, D., & Terpenny, J. (2020). Decision-level data fusion in quality control and predictive maintenance. *IEEE Transactions on Automation Science and Engineering*, *18*(1), 184-194.

[14] Fikardos, M., Lepenioti, K., Bousdekis, A., Bosani, E., Apostolou, D., & Mentzas, G. (2022, September). An Automated Machine Learning Framework for Predictive Analytics in Quality Control. In Advances in Production Management Systems. Smart Manufacturing and Logistics Systems: Turning Ideas into Action: IFIP WG 5.7 International Conference, APMS 2022, Gyeongju, South Korea, September 25–29, 2022, Proceedings, Part I (pp. 19-26). Cham: Springer Nature Switzerland.

[15] Ammar, M., Haleem, A., Javaid, M., Walia, R., & Bahl, S. (2021). Improving material quality management and manufacturing organizations system through Industry 4.0 technologies. *Materials Today: Proceedings*, *45*, 5089-5096.

[16] Belov, A. V., Melekhova, E. A., & Vorontsova, T. (2022, September). Predictive Analytics Approach for Steel Billets Quality Control System. In 2022 International Conference on Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS) (pp. 219-223). IEEE.

[17] Voisin, A., Laloix, T., Iung, B., & Romagne, E. (2018). Predictive maintenance and part quality control from joint product-process-machine requirements: Application to a machine tool. Procedia Manufacturing, 16, 147-154.

[18] Cui, P. H., Wang, J. Q., & Li, Y. (2022). Data-driven modelling, analysis and improvement of multistage production systems with predictive maintenance and product quality. International Journal of Production Research, 60(22), 6848-6865.

[19] Zonnenshain, A., & Kenett, R. S. (2020). Quality 4.0—the challenging future of quality engineering. Quality Engineering, 32(4), 614-626.

[20] Lade, P., Ghosh, R., & Srinivasan, S. (2017). Manufacturing analytics and industrial internet of things. IEEE Intelligent Systems, 32(3), 74-79.

[21] Colosimo, B. M., Huang, Q., Dasgupta, T., & Tsung, F. (2018). Opportunities and challenges of quality engineering for additive manufacturing. Journal of Quality Technology, 50(3), 233-252.

[22] Pulikottil, T., Estrada-Jimenez, L. A., Nikghadam-Hojjati, S., & Barata, J. (2021). Predictive manufacturing: enabling technologies, frameworks and applications. In Technological Innovation for Applied AI Systems: 12th IFIP WG 5.5/SOCOLNET Advanced Doctoral Conference on Computing, Electrical and Industrial Systems, DoCEIS 2021, Costa de Caparica, Portugal, July 7–9, 2021, Proceedings 12 (pp. 51-61). Springer International Publishing.

[23] Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2021). Significance of Quality 4.0 towards comprehensive enhancement in manufacturing sector. Sensors International, 2, 100109.

[24] Kumar, M., Shenbagaraman, V. M., Shaw, R. N., & Ghosh, A. (2021). Predictive data analysis for energy management of a smart factory leading to sustainability. In Innovations in Electrical and Electronic Engineering: Proceedings of ICEEE 2020 (pp. 765-773). Springer Singapore.

[25] Alok, S., Kabra, A., & Mudam, A. (2018). Predicting the behavioural intention to adopt lean practices: an empirical study in the manufacturing industry. International Journal of Services and Operations Management, 29(4), 557-578.

[26] S. Oh, J. Cha, D. Kim and J. Jeong, "Quality Inspection of Casting Product Using CAE and CNN," 2020 4th International Conference on Imaging, Signal Processing and Communications (ICISPC), 2020, pp. 34-38, doi: 10.1109/ICISPC51671.2020.00014.

[27] Aashish Arora and Rajeev Gupta. “A Comparative Study on Application of Artificial Intelligence for Quality Assurance in Manufacturing”, IEEE International conference on Inventive Research in Computing Applications. 2022.

[28] https://www.kaggle.com/datasets/fantacher/neu-metal-surface-defects-data

[29] Alves, J.B., Marques, B., Dias, P. et al. Using augmented reality for industrial quality assurance: a shop floor user study. Int J Adv Manuf Technol 115, 105–116 (2021). https://doi.org/10.1007/s00170-021-07049-8

[30] Jacqueline Schmitt, Jochen Bönig, Thorbjörn Borggräfe, Gunter Beitinger, Jochen Deuse, Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing, Advanced Engineering Informatics, Volume 45, 2020, 101101, ISSN 1474-0346, https://doi.org/10.1016/j.aei.2020.101101

[31] Hasibur Rahman, Md. Omar Faruq, Talha Bin Abdul Hai, Wahidur Rahman, Muhammad Minoar Hossain, Mahbubul Hasan, Shafiqul Islam, Md. Moinuddin, Md. Tarequl Islam, Mir Mohammad Azad, IoT enabled mushroom farm automation with Machine Learning to classify toxic mushrooms in Bangladesh, Journal of Agriculture and Food Research, Volume 7, 2022, 100267, ISSN 2666-1543, https://doi.org/10.1016/j.jafr.2021.100267.

[32] Nikolić, Filip, Ivan Štajduhar, and Marko Čanađija. "Casting defects detection in aluminum alloys using deep learning: A classification approach." International Journal of Metalcasting 17.1 (2023): 386-398.

[33] Parlak, İsmail Enes, and Erdal Emel. "Deep learning-based detection of aluminum casting defects and their types." Engineering Applications of Artificial Intelligence 118 (2023): 105636.