NextGen Scientific Publication
American International Journal of Computer Science and Technology
Volume 8 Issue 3, Pg. No. 50-60, AIJCST-V8I3P105, 2026
https://doi.org/10.63282/3117-5481/AIJCST-V8I3P105

Original Article

Machine Learning-Driven Automated Quality Inspection of Steel Plates in Industrial Production Lines

*** Aravindh Balan**
Freelance Post-doctoral scholar, Project Manager, Inline hydraulics GmbH, Germany.

Abstract:

Steel is a vital commodity in many different industries and the building trades. It's crucial to have high-quality steel plates to avoid any potential problems. An essential step in detecting and classifying surface defects in rolled metal is the pre-distribution quality evaluation. This study provides a CNN+BiLSTM hybrid neural network system that can automatically detect and categorize surface flaws on steel plates. The Faulty Steel Plates dataset forms basis of the technique proposed in this paper. Cleaning dataset, normalizing it using Min-Max, balancing it using SMOTE, and stratifying it are all data preprocessing procedures that are utilized to improve the dataset's quality and classification performance. In order to improve learning, CNN component records most crucial spatial features and the BiLSTM component records the sequential dependency. The performance of the proposed model is compared with traditional machine learning models such as Decision Tree (DT), Random Forest (RF), and K-Nearest Neighbor (KNN). Experimental findings showed that the suggested model (CNN+BiLSTM) outperformed existing models with an F1-score (F1), recall (rec), accuracy (acc), and precision (prec) of 99.7%. The findings prove that the suggested technique is an effective and dependable way to automatically detect surface defects in steel plates in an industrial production environment.

Keywords:

Artificial Intelligence, Deep Learning, Quality Control, Visual Inspection, Smart Manufacturing, Defect Detection.

Article History:

Received: 22.03.2026

Revised: 25.04.2026

Accepted: 02.05.2026

Published: 10.05.2026

1. Introduction

The industrial technology has significantly transformed modern manufacturing systems by improving productivity, efficiency, and product quality [1]. Among various industrial materials, steel remains one of the most widely used due to its strength, durability, and adaptability in applications including construction, automotive manufacturing, machinery, and pipelines [2]. Steel sheets are commonly produced by rolling, in which metal passes between rotating rollers to achieve the required thickness and shape [3]. There may be defects in the manufacturing process, however, because of temperature fluctuations, mechanical stress or process instability, which have a direct impact on product quality and industrial reliability [4].

The main steel surface inspection methods were manual visual inspection and the traditional non-destructive testing methods [5]. Stroboscopic lighting systems were commonly used in industrial environments for detecting defects on moving steel plates [6]. These methods have been popular but had a number of disadvantages such as high labor burden, visual fatigue, inaccuracy in detection and false detection rates [7]. In order to overcome these limitations, industries went step by step to the automated inspection systems, which were based on computer vision and intelligent monitoring technologies for more accurate defect detection.

In the industrial sectors, smart manufacturing and quality assurance have seen tremendous advances with development of Artificial Intelligence (AI), Internet of Things (IoT) and automation technologies [8]. By using AI-driven systems, production information can be analyzed quickly, resulting in lower operational expenses, fewer human errors, and better decision-making processes [9]. When considering product quality, quality control is becoming increasingly important in the competitive industrial environment to satisfy customers, reduce costs and enhance production efficiency [10]. But there are still some challenges, such as model interpretability and adaptability to industry, that are significant issues for wide application of intelligent inspection technologies [11].

In recent years, Machine Learning (ML) and Deep Learning (DL) have emerged as effective solutions for automated steel plate quality inspection and defect classification [12]. The ML algorithms can be trained with production data to understand the patterns of defects, and accurately classify steel surface defects [13][14] and the DL-based intelligent framework can automatically extract complex features of the defect images with high precision. The combination of ML and DL methods and real-time industrial monitoring systems can enable industries to achieve intelligent fault detection and improve production efficiency.

1.1. Motivation and Contribution

The motivation for this study is the high demand for high-quality steel products in modern industries and the development of accurate and automatic quality inspection system. In traditional manual inspection, the inspection process is time-consuming and labor-intensive and easily has human errors, especially in high-speed production. With recent advances in ML and DL, efficient solutions for detecting and classifying steel surface defects with enhanced accuracy and reliability are now available. Thus, this study's purpose is to determine an intelligent automated inspection framework to improve production efficiency, decrease defective products and increase industrial quality assurance. In the following, this study is shown to have a number of important contributions:

- Utilized the Faulty Steel Plates dataset for comprehensive steel surface defect analysis.
- Applied effective preprocessing techniques including data cleaning, Min-Max normalization, SMOTE balancing, and stratified data splitting to improve dataset quality.
- Proposed a hybrid CNN+BiLSTM DL model for accurate steel plate surface defect classification and quality assessment.
- Demonstrated effectiveness of proposed model for automated and reliable steel plate quality assessment in industrial production environments.
- Several performance metrics are used to evaluate proposed model, including rec, acc, prec, and F1, in order to provide a thorough evaluation of its classification performance.

The justification for this research lies in the increasing need for efficient, automated steel plate surface defect detection systems in industrial production environments, where traditional inspection methods often suffer from limited accuracy and high dependence on manual analysis. Existing ML approaches struggle to capture complex defect characteristics and sequential relationships among features. This study's unique contribution is a CNN+BiLSTM hybrid model that combines CNN's spatio-temporal feature extraction capabilities with BiLSTM's bidirectional sequential learning capabilities. This hybrid framework provides a more robust, intelligent, and reliable solution for steel plate quality assessment and defect classification.

1.2. Organization of the Paper

The paper is organized as follows: Section II provides an overview of relevant work; Section III details dataset, preprocessing, and proposed model, Section IV discusses experimental results and comparative analysis, and Section V concludes study with suggestions for future research.

2. Literature Review

ML and DL models used to assess steel plate quality and detect surface defects are reviewed in this section.

Y. Zhu et al. (2026) focused on specific defect categories, reducing false positives in challenging texture-including defects. Experiments on public NEU-DET dataset show that improved model achieves an mAP50 of 92.9%, a 7.2% gain over baseline. Notably, detection performance for small and blurry crazing category improves by 14.8% in mAP50, validating the effectiveness of the proposed approach [15]. B. Jiang et al. (2025) aimed to raise the quality of the image data that was produced. Lastly, the DCNN is trained using the larger dataset. The trained model achieves a recognition accuracy of 96.16% on the test set. The proposed approach has strong technical applicability for identifying surface imperfections in steel plates, according to experimental results using real industrial data, in contrast to human inspection's high error rate and the difficulties of obtaining and labeling datasets [16]. Xue et al. (2025) proposed to integrate hardware systems with industrial cameras and computer servers, while Python programming software is utilized for system development. The system has a recognition running time of less than 0.5 seconds, an offset recognition accuracy of 95 %, good precision, meets industrial production needs, reduces labor intensity, and improves production efficiency [17]. V. Vasan et al. (2024) found that the suggested model achieves an impressive overall accuracy of 96.39% when it comes to identifying and categorizing surface faults in steel. Aside from describing the vision transformer architecture, the paper also compares current model's performance to that of alternative methodologies proposed for use in literature and provides descriptive insights into both [18]. A. Feyzioğlu and Y. S. Taspınar (2023) There are four distinct ML techniques that may be utilized to identify flaws when categorization procedures are implemented. Iterative classification procedures make use of algorithms like LR, DT, SVM, and RF. With an RF model of 79.44%, achieve the best possible classification accuracy. When investigating how different dataset properties affect classification accuracy, correlation analysis is a useful tool. According to experts, the suggested models have adequate classification accuracy for this difficult problem, but they believe it could be improved [19]. X. Feng et al. (2022) Based on real-life industrial background, the goal was to use ResNet152 and other deep learning methods to find flaws in the hot rolling flattening step. The experimental results on the XL Data-CLS dataset show that the method in this paper achieves 97.64% classification accuracy, which meets the requirements of practical detection[20].

Based on recent research, Table I outlines the proposed models, key findings and challenges faced in Steel Plate Quality assessment.

Table 1. Recent Studies on Steel Plate Quality Assessment Using Machine Learning Technqiues

Author	Proposed Work	Results	Key Findings	Limitations & Future Work
Zhu, Liu and Li (2026)	Improved defect detection model focusing on challenging texture-including steel defects using the NEU-DET dataset.	Achieved 92.9% mAP50 with a 7.2% improvement over baseline; crazing defect detection improved by 14.8%.	Enhanced detection capability for small and blurry defect categories with reduced false positives.	Further improvement is needed for real-time deployment and testing on larger industrial datasets.
Jiang et al. (2025)	Proposed a DCNN-based defect recognition framework with extended image dataset generation for industrial steel defect inspection.	Obtained 96.16% recognition accuracy on the test dataset.	Demonstrated effective dataset augmentation and improved industrial applicability for automated inspection.	Future work may focus on reducing computational complexity and improving dataset diversity.
Xue et al. (2025)	Developed a hardware-integrated industrial vision system using cameras, servers, and Python-based software for steel defect detection.	Achieved 95% recognition accuracy with processing time below 0.5 seconds.	The system met industrial production requirements with improved efficiency and reduced labor intensity.	Future studies can enhance scalability and optimize performance for complex production environments.
Vasan et al. (2024)	Developed a model utilizing Vision Transformer to identify and categorize surface flaws in steel.	Achieved an overall accuracy of 96.39%.	Demonstrated effectiveness of transformer architectures for industrial surface defect analysis.	Further research can improve lightweight implementation and real-time industrial deployment.
Feyzioğlu and Taspınar (2023)	Applied LR, DT, SVM, and RF algorithms for steel defect classification with correlation analysis.	RF achieved the highest accuracy of 79.44%.	Feature correlation analysis helped evaluate the impact of dataset attributes on classification	The obtained accuracy is satisfactory but requires enhancement through advanced hybrid or deep learning models.

			performance.	
Feng, Gao and Luo (2022)	Utilized ResNet152 and deep learning approaches for defect detection during the hot rolling flattening stage.	Achieved 97.64% classification accuracy on the XLData-CLS dataset.	The proposed deep learning model satisfied practical industrial detection requirements.	Future work may address computational efficiency and adaptability to varying industrial conditions.

Research gaps: There are several ML and DL models proposed for defect detection on steel plate surfaces, but many models are still unable to satisfy high classification accuracy and efficient feature extraction, and suitable for industrial application and real-time process. While traditional ML algorithms are successful for simple spatial and sequential defects, certain DL models are complex in terms of computation and require extensive training on all types of manufacturing data. Moreover, there has been little work on the simultaneous integration of spatial and temporal learning ability in a single system for defect classification. Hence, the need of the hour is for an efficient hybrid DL model that can offer a reliable and accurate assessment of the quality of steel plates in the industrial environment in an automated way.

3. Research Methodology

The suggested approach in identifying steel plate quality is based on steel plate surface defect detection method using Faulty Steel Plates Dataset, which is a collection of steel plates with various defects on the surface. The data is preprocessed using techniques such as data cleaning, normalization, balancing, and stratified data splitting to enhance data quality and reliability for the model. A hybrid CNN–Bi-LSTM network is designed. Standard classification metrics and model performance is assessed using confusion matrix analysis, as seen in Fig. 1.

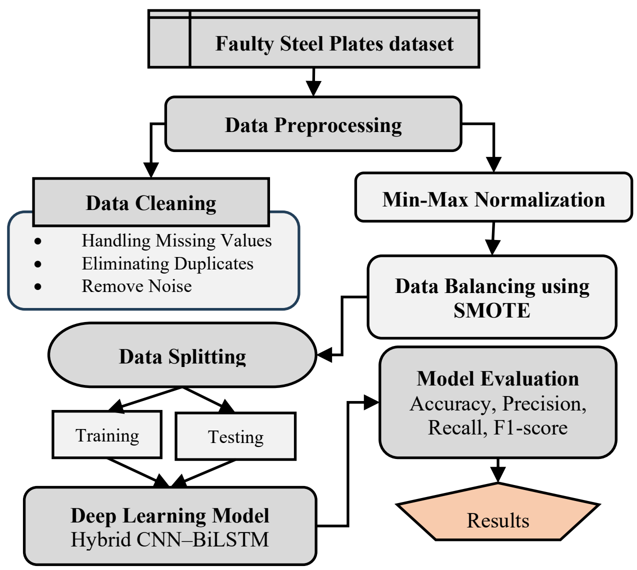

Figure 1.Proposed flowchart for Steel Plate Quality Assessment using machine learning

This section provides details about each step of the proposed methodology:

3.1. Data Gathering and Analysis

This study uses the Faulty Steel Plates Dataset [21], which contains 1,941 samples with 34 attributes for steel surface defect classification. The dataset includes 27 feature measurements and seven defect classes representing different types of surface faults in stainless steel plates. Its purpose is to teach and test automatic fault detection models that use DL and ML. Distribution and feature correlations are examined using data visualizations etc., are given below:

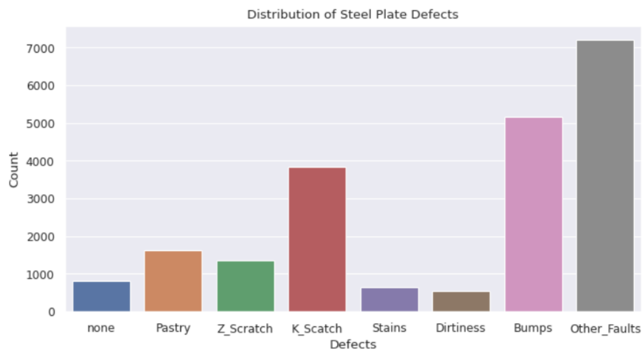

Figure 2. Distribution By Classes In The Dataset

Fig. 2 depicts the class-wise distribution of the data set which reveals some imbalance among the different classes of defects. There are many more samples in the “Other Faults” class, and many fewer in classes including “Stains” and Dirtiness. The imbalance is indicative of the need for balancing techniques for effective and fair training of the model.

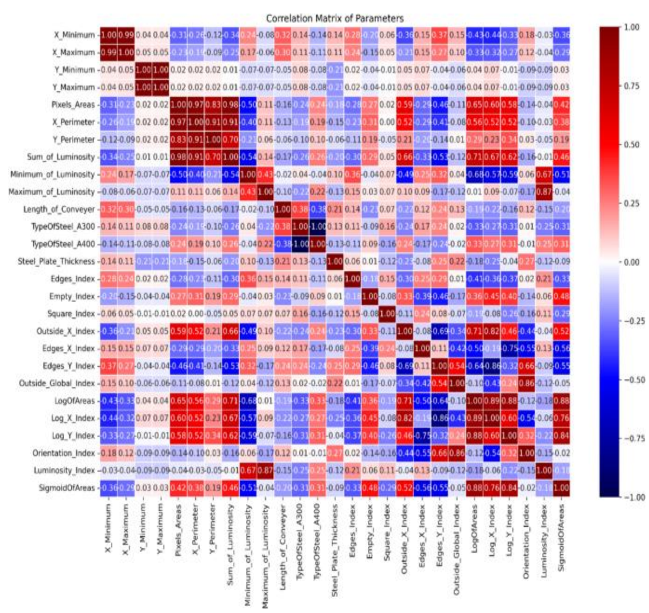

Figure 3.Correlation Matrix Heatmap On Faulty Steel Plates Dataset For Steel Plate Quality Assessment

Fig. 3 presents the correlation matrix heatmap for the faulty steel plates dataset illustrates the relationships among various extracted features used for steel plate quality assessment. It shows very clearly that there are positive and negative correlations between the parameters, thus there may be some redundancy and interdependence between the variables.

3.2. Data Pre-Processing

The Faulty Steel Plates dataset is used for data preparation, which involves concatenation, cleansing, and feature engineering. The data pre-processing steps undertaken involved handling missing values, removal of duplicates, removal of redundant data and noise, data labelling and normalization. The following are the main preprocessing procedures:

3.3. Data Cleaning

Data preprocessing addresses missing data by identifying and eliminating missing data points and missing class labels to ensure the data set's consistency and reliability. Duplicate and redundant entries are eliminated, and the selected columns to prevent repetition. Additionally, suitable preprocessing techniques accomplish noise reduction, guaranteeing data quality and overall machine learning model performance.

3.4. Feature Scaling with Min–Max Normalization

The min-max method, which scales from 0 to 1, is used to normalize the records. This to make the classifiers work better and lessen the effect of the outliers. When it comes to normalizing, the mathematical Equation (1) is utilized.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

where X is feature's initial value, $X'$is its normalized value, X_{min} is its minimum value, and X_{max} is its maximum value.

3.5. Data balancing using Synthetic Minority Over-Sampling Technique (SMOTE)

Data balancing is process of addressing imbalances in data sets, wherein one class has significantly fewer samples compared to the other. The precision and effectiveness of ML models may be impacted by this disparity. In order to enhance the class distribution and increase number of minority samples in dataset, this study uses SMOTE synthetic sample generation approach. In order for SMOTE to create new samples, it calculates the distance in geometric units between each minority class sample and the closest samples of the same class.

3.6. Data Splitting

Training and test sets make for 80:20 of the dataset. The dataset is divided 80:20 between training and testing. This method allows the classes to have same proportional class distribution as the original data. It, therefore, contributes to maintain data balance and to enhancing reliability of model evaluation.

3.7. Proposed Model

The hybrid CNN–BiLSTM Model for Steel Plate Quality Assessment in Industrial Production Lines is a DL-based hybrid model that is suggested in this research.

3.7.1. *Convolutional Neural Network (CNN) Model*

CNNs have been widely adopted for classification problems because of their high performance (accuracy). The basic architecture of a CNN consists of the input, convolution, activation, pooling, fully connected, and output layers. The original image data is entered into the CNN's input layer. For each pixel in the image, there is an input neuron. Therefore, the CNN is based on use of convolutional layers, which are basic elements of the CNN. These layers employ convolution processes to extract local information from input image. Several filters, also known as kernels, are slid over the image during convolution process to calculate dot products with regional patches. Equation (2) yields another image matrix y, which is the convolution of image X with filter H. When i and j are outside of the range, X_{ij} is regarded as zero.

$$y_{ij} = \sum_{a=1}^{n} \sum_{b}^{n} H_{ab} x_{(i+a)(j+b)} (2)$$

The filter's sliding motion across the input image and the exact dimensions of the output image are controlled by two common parameters, padding and stride, in addition to filter size. One output of CNN is a feature map that draws attention to certain patterns, such as textures or edges. By first removing negative values with a ReLU activation and then pooling (max or average) to downsample the feature map, can reduce dimensionality without losing any important information. Equation (3, 4) gives the output's dimensions following a pooling layer for a feature map with dimensions $q_h \times q_w \times q_c$.

$$ya_{ij} = f\,(y_{ij}) = \max\,(0, y_{ij}) \ (3)$$

$$r_h \times r_w \times r_c = \left(q_p - p + 1\right)/s \times (q_w - p + 1)/sxq_c \quad (4)$$

The feature map's height $\mathrm{q_h}$, feature map width $\mathrm{q_w}$, feature map number of channels $\mathrm{q_c}$, pooling filter size (p), stride of pooling layer (s), downsampled feature height $\mathrm{r_h}$, downsampled feature width $\mathrm{r_w}$, and downsampled feature number $\mathrm{r_c}$ are all variables in this context. Activation functions, pooling layers, convolution layers, and finally a fully connected (FC) layer are used for classification. The FC layer establishes connections with all neurons in the layer below it. Class probability is provided by the final layer, which is composed of SoftMax activation units.

3.7.2. *Bidirectional Long Short-Term Memory (Bi-LSTM) Model*

In Bi-LSTM, data is moved from the past to the future by the forward hidden layer and vice versa by the backward hidden layer. When compared to traditional LSTM, Bi-LSTM demonstrates superior data representation capabilities in the context of deep learning architectures. What follows is an explanation of the Bi-LSTM output Equations (5) to (7):

$$h^f{}_t = LSTM(x_t, h^f{}_{t-1}) \quad (5)$$
$$h^b{}_t = LSTM(x_t, h^b{}_{t-1}) \quad (6)$$
$$y_t = W_o h_t + b_o \quad (7)$$

The weights from forward layer to output layer are represented by $W^f{}_{hy}$, weights from backward layer to output layer are represented by $W^b{}_{hy}$, and bias vector of output layer is b_o. $h^f{}_t$ and $h^b{}_t$ are the components that make up h_t. Bi-LSTM uses both sets of data concurrently to learn at time t by combining information from past and future.

3.7.3. *Proposed hybrid CNN–BiLSTM Model*

The CNN–BiLSTM model combines temporal sequence learning skills of Bi-LSTM with the spatial feature extraction capabilities of CNN. High-level spatial characteristics are extracted from input data, such as images or time-series representations, by the CNN component of this hybrid architecture, whereas Bi-LSTM component captures long-term interdependence in both forward and backward directions. This combination enables model to learn more comprehensive and discriminative representations of the data. The convolutional process between an input image X and a filter H is defined in Equation (8) as:

$$y_{ij} = \sum_{a=1}^{n} \sum_{b=1}^{n} H_{ab} X_{(i+a)(j+b)} \quad (8)$$

In this context, y_{ij} is convolved feature map, H_{ab} is convolution kernel, and $X_{(i+a)(j+b)}$ is local region in input image that is being convolved. Following this, in order to introduce non-linearity and maintain only positive values, a ReLU activation function is utilized. Pooling is then used to minimize the spatial dimensions while preserving dominating characteristics. Equation (9) displays the dimensions of the output feature map following pooling:

$$r_h \times r_w \times r_c = \frac{(q_h - p + 1)}{s} \times \frac{(q_w - p + 1)}{s} \times q_c \quad (9)$$

In the feature map, q_h,q_w,q_c represent the number of height, width, and channel, respectively, of the feature map, and p and s represent the size of the pooling filter and the stride, respectively. The CNN flattened feature maps are then fed into the Bi-LSTM layer for sequential learning.

These features are processed in both time directions using a Bi-LSTM network. Dependency from past to future is captured by forward layer, and dependency from the future to past is captured by backward layer. The model can remember the context from the entire input sequence, to this two-way processing. To get the best learning results, the proposed hybrid CNN+BiLSTM is run with a learning rate of 0.001, a batch size of 32, and 100 training epochs. The BiLSTM block consisted of 128 hidden units to learn the bidirectional sequence, whereas the CNN block consisted of 64 3x3 convolutional filters with ReLU activation. SoftMax activation in output layer and categorical cross-entropy as loss function are used to train Adam optimizer model.

3.7.4. *Evaluation Metrics*

Various classification parameters are used to assess proposed model. To demonstrate the number of right and wrong guesses for each class, a confusion matrix is first generated. True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) are extracted from this matrix and utilized to compute pertinent assessment metrics, such as acc, prec, rec, and F1:

Accuracy: A key performance indicator for evaluating a classification model's overall effectiveness. It is the proportion of examples in dataset that were properly predicted to all instances. It is given as Equation (10)-

$$Accuracy = \frac{\text{TP+TN}}{\text{TP+Fp+TN+FN}} \quad (10)$$

Precision: Precision assesses how accurate classifier's affirmative prediction was. Precision shows percentage of truly affirmative cases. How good a classifier is in predicting the positive classes is expressed as Equation (11)-

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \quad (11)$$

Recall: Recall is sometimes referred to as TPR or sensitivity. It is ratio of all TP predictions to dataset's actual positive occurrences. Recall measures a model's ability to locate all relevant examples. In mathematical form, it is given as Equation (12)-

$$Recall = \frac{\text{TP}}{TP+FN} \qquad (12)$$

F1 score: The F1 score is the harmonic mean of accuracy and recall. It provides a single statistic that balances accuracy and recall. When there is an unequal distribution of classes, it is very useful. Mathematically, it is given as Equation (13)-

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (13)$$

4. Results and Discussion

An NVIDIA Titan graphics processing unit (GPU), two 10-core Intel Xeon E5-2600 v4 CPUs, and 512 GB of RAM make up the graphics workstation used for all tests. The programming environment is Windows 11, and the learning framework is Keras.

4.1. Performance Evaluation

In this section, proposed hybrid CNN+BiLSTM model is described, and performance is evaluated during the training and testing process with Faulty Steel Plates dataset. Important performance measures for model to assess are included in Table II, including acc, prec, rec, and F1. The results demonstrate high classification accuracy: the model achieves 99.7% acc, with prec, rec, and F1 all at 99.7%, indicating that it is highly reliable, exhibits well-balanced performance, and can detect defects with high accuracy in classifying steel plates.

Table 2.Classification Results of the Proposed Hybrid CNN+Bilstm Model for Faulty Steel Plates Dataset

Matrix	Testing	Training
Accuracy	99.7	1.0
Precision	99.7	1.0
Recall	99.7	1.0
F1-score	99.7	1.0

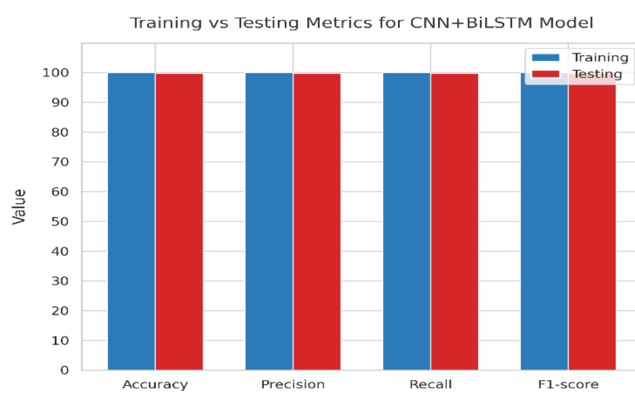

Figure 4.Train and Test performance comparison of proposed hybrid CNN+BiLSTM Model for Steel Plate Quality Assessment

The grouped bar chart of CNN+BiLSTM model's acc, prec, rec and F1 is shown in Fig. 4. The model achieves 100% across all training metrics. It has been consistently evaluated as 99.7% accurate in all evaluation measures, and has good generalization ability, which is desirable for testing.

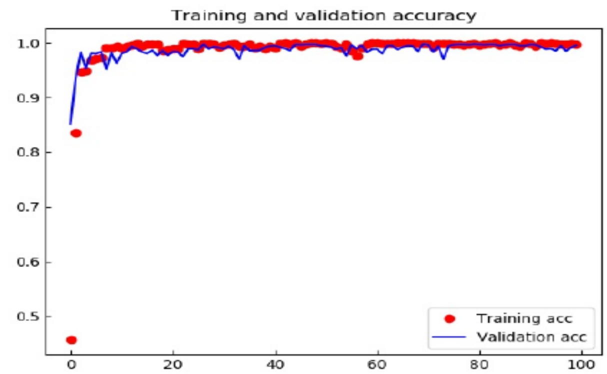

Figure 5. Acc Curve For The Hybrid CNN+Bilstm Model

The CNN+BiLSTM model's training and validation accuracy curves are shown in Fig. 5. The graph shows that in early epochs, the accuracy of both training and validation data rapidly increases to a high value close to 1.0. This means model performs well on learning task, provides low overfitting and has a good generalization capacity.

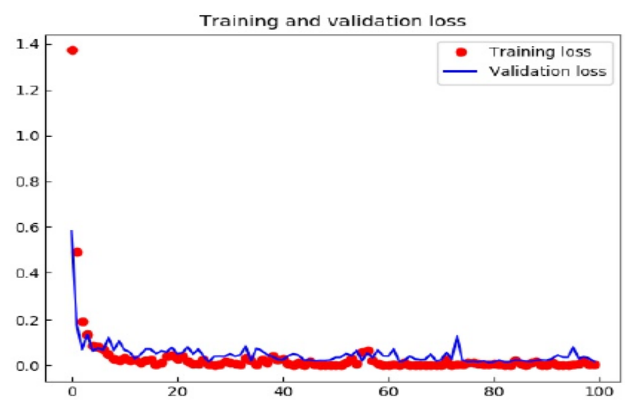

Figure 6.Loss Curve For Hybrid CNN+Bilstm Model

The CNN+BiLSTM model's training and validation loss curves after 100 epochs are displayed in Fig. 6. The training loss initially increases steeply and then quickly approaches zero, suggesting that training process is successful and model is converging. The validation loss also shows a similar trend to the training loss, albeit with some minor variations, confirming that model generalizes well and is not heavily overfitting.

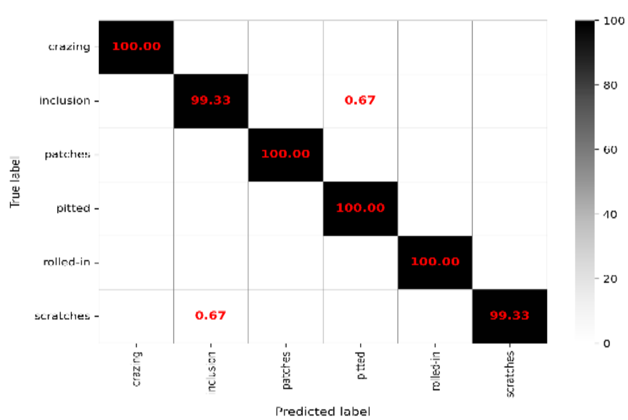

Figure 7.Confusion Matrix For CNN+Bilstm Model

Fig. 7 show normalized confusion matrix evaluates CNN+BiLSTM model across six surface defect classes, showing near-perfect classification performance. The majority of classes including crazing, patches, pitted and rolled-in have a 100% accuracy rate on the diagonal. There is very little confusion between inclusion and scratches with only a very small 0.67% misclassification between those two classes.

4.2. Comparative Analysis

The proposed hybrid CNN+BiLSTM is compared with previously mentioned machine learning models, as shown in Table III, to evaluate its effectiveness. The Random Forest and KNN models produced higher accuracy of 90.14% and 91.7%, respectively, while the Decision Tree model produced the lowest accuracy of 77.7%. The proposed hybrid CNN + BiLSTM model outperformed existing models in terms of performance, reliability, and efficiency in classification of steel plate defects, achieving highest acc of 99.7% with high prec of 99.3%, rec of 99.7%, and F1 of 99.3%.

Table 3.Comparison of Different Machine Learning and Deep Learning Models For Steel Plate Quality Assessment

Model	Accuracy	Precision	Recall	F1-score
DT [19]	77.7	81	74	77
RF [22]	90.14	84.75	75.89	77.61
KNN [23]	91.7	86	87.1	87.1
Proposed hybrid CNN+BiLSTM Model	99.7	99.7	99.7	99.7

5. Conclusion and Future Study

Steel plate defect techniques have been widely used in industrial applications to assess the quality of steel plate surfaces. Traditional non-destructive testing methods and then manual visual inspection are used for defect inspection, this takes a long time and is prone to human error. The accuracy and degree of automation in defect identification have significantly increased because to advancements in AI technology, machine vision, and DL techniques. This paper proposes a CNN+BiLSTM hybrid model to evaluate the steel plate quality in Faulty Steel Plates. The proposed model's accuracy is 99.7%, which is higher than the traditional algorithms including DT, RF and KNN with an accuracy of 77.7%, 90.14% and 91.7% respectively. CNN and BiLSTM presented effective and robust spatial and sequential features performance in classification, with better performance. Future work can focus on real-time industrial deployment and evaluation on larger datasets to improve robustness of proposed system. Advanced models such as ResNet, EfficientNet, and attention-based hybrid models can be incorporated to improve further the defect detection and classification accuracy.

References

[1] A. Krommuang and O. Suwunnamek, “Internet of Things (IoT) Application for Management in Automotive Parts Manufacturing,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 4, 2022, doi: 10.14569/IJACSA.2022.0130474.

[2] Y. A. Purmala, “Implementation of machine learning to increase productivity in the manufacturing industry: a literature review.,” *Oper. Excell. J. Appl. Ind. Eng.*, vol. 13, no. 2, p. 267, Jul. 2021, doi: 10.22441/oe.2021.v13.i2.026.

[3] T. Kulkarni, B. Toksha, S. Shirsath, S. Pankade, and A. T. Autee, “Construction and Praxis of Six Sigma DMAIC for Bearing Manufacturing Process,” *Mater. Today Proc.*, vol. 72, pp. 1426–1433, 2023, doi: 10.1016/j.matpr.2022.09.342.

[4] D. S. Das, V. Challagulla, D. P. Balaramesh, D. A. M. A. Mohan, B. V. Jyothi, and D. S. Saravanan, “Deep Learning for Corrosion Monitoring Virtual Sensor and Predictive Modelling Approaches in Industrial Water Pipeline,” *The Bioscan*, vol. 21, no. 1, pp. 35–45, Jan. 2026, doi: 10.63001/tbs.2026.v21.i01.pp35-45.

[5] T. H. Febriana and H. Hasbullah, “Analysis and Defect Improvement Using FTA, FMEA, and MLR Through DMAIC Phase: Case Study in Mixing Process Tire Manufacturing Industry,” *J. Eur. des Systèmes Autom.*, vol. 54, no. 5, pp. 721–731, Oct. 2021, doi: 10.18280/jesa.540507.

[6] S. Farahmand-Tabar and P. Ashtari, “Simultaneous size and topology optimization of 3D outrigger-braced tall buildings with inclined belt truss using genetic algorithm,” *Struct. Des. Tall Spec. Build.*, vol. 29, no. 13, Sep. 2020, doi: 10.1002/tal.1776.

[7] P. Ashtari, R. Karami, and S. Farahmand-Tabar, “Optimum geometrical pattern and design of real-size diagrid structures using accelerated fuzzy-genetic algorithm with bilinear membership function,” *Appl. Soft Comput.*, vol. 110, p. 107646, Oct. 2021, doi: 10.1016/j.asoc.2021.107646.

[8] S. Farahmand-Tabar and M. Babaei, “Memory-assisted adaptive multi-verse optimizer and its application in structural shape and size optimization,” *Soft Comput.*, vol. 27, no. 16, pp. 11505–11527, Aug. 2023, doi: 10.1007/s00500-023-08349-9.

[9] S. Kilaru, “Automated ETL Intelligence: Metadata-Orchestrated Framework with Rule-Based Heuristics for Monitoring and Reporting,” *Int. J. Inf. Electron. Eng.*, vol. 3, no. 6, p. 14, Aug. 2014, doi: 10.48047/ijiee.2013.3.6.9.

[10] N.-Z. Chen, Z. Zhao, and L. Lin, “A hybrid deep learning method for AE source localization for heterostructure of wind turbine blades,” *Mar.*

Struct., vol. 94, p. 103562, Mar. 2024, doi: 10.1016/j.marstruc.2023.103562.
[11] A. K. Padhy, N. Seshagiri, V. Soni, A. K. Elengovan, G. Babu Thokala, and M. Kumari, “AI-Enabled Autonomous Data Preprocessing: A Scalable Architecture for Intelligent Machine Learning Pipeline Management,” in *2026 6th International Conference on Image Processing and Capsule Networks (ICIPCN)*, IEEE, Jan. 2026, pp. 1318–1323. doi: 10.1109/ICIPCN67432.2026.11438503.
[12] Y. Liu, K. Xu, and J. Xu, “Periodic Surface Defect Detection in Steel Plates Based on Deep Learning,” *Appl. Sci.*, vol. 9, no. 15, p. 3127, Aug. 2019, doi: 10.3390/app9153127.
[13] R. Tang and K. Mao, “An Improved GANs Model for Steel Plate Defect Detection,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 790, no. 1, p. 012110, Mar. 2020, doi: 10.1088/1757-899X/790/1/012110.
[14] D. Soukup and R. Huber-Mörk, “Convolutional Neural Networks for Steel Surface Defect Detection from Photometric Stereo Images,” 2014, pp. 668–677. doi: 10.1007/978-3-319-14249-4_64.
[15] Y. Zhu, Y. Liu, and H. Li, “An Improved RT-DETR Method for Surface Defect Detection of Small Targets on Steel Plates,” in *Proceedings of the 2026 International Conference on Artificial Intelligence and Control*, New York, NY, USA, NY, USA: ACM, Feb. 2026, pp. 128–134. doi: 10.1145/3807246.3807268.
[16] B. Jiang, P. Zhou, X. Sun, and T. Chai, “Intelligent recognition of steel plate surface defect based on deep convolutional GAN,” *Neural Comput. Appl.*, vol. 37, no. 14, pp. 8331–8345, May 2025, doi: 10.1007/s00521-025-11007-w.
[17] R. Xue *et al.*, “Research on Online Detection Method for Lateral Deviation of Steel Strip in Steel Production Line Based on Machine Vision,” in *2025 IEEE International Conference on Pattern Recognition, Machine Vision and Artificial Intelligence (PRMVAI)*, IEEE, Jun. 2025, pp. 1–6. doi: 10.1109/PRMVAI65741.2025.11108523.
[18] V. Vasan, N. V. Sridharan, S. Vaithiyanathan, and M. Aghaei, “Detection and classification of surface defects on hot-rolled steel using vision transformers,” *Heliyon*, vol. 10, no. 19, p. e38498, Oct. 2024, doi: 10.1016/j.heliyon.2024.e38498.
[19] A. Feyzioğlu and Y. S. Taspınar, “Detection of Defects in Rolled Stainless Steel Plates by Machine Learning Models,” *Int. J. Appl. Math. Electron. Comput.*, vol. 11, no. 1, pp. 37–43, Mar. 2023, doi: 10.18100/ijamec.1253191.
[20] X. Feng, X. Gao, and L. Luo, “Research on the application of deep learning based surface defect detection and treatment method for hot rolled strip steel,” in *2022 41st Chinese Control Conference (CCC)*, IEEE, Jul. 2022, pp. 7361–7366. doi: 10.23919/CCC55666.2022.9902418.
[21] C. Crawford, “Faulty Steel Plates,” Kaggle. [Online]. Available: https://www.kaggle.com/datasets/uciml/faulty-steel-plates
[22] H. Vaidya, K. V. Prasad, R. S, K. S. K, and S. Y, “Enhancing Steel Manufacturing Quality Control: A Deep Learning Approach for Rapid and Accurate Surface Defect Detection Using Pre-Trained Convolutional Neural Network Features and Supervised Classifiers,” *Int. Acad. Phys. Sci.*, vol. 28, no. 1, pp. 1–18, 2024, doi: 1061294/jiaps2024.2811.
[23] B. Tasar, “Comparison Analysis of Machine Learning Algorithms for Steel Plate Fault Detection,” *Düzce Univ. J. Sci. Technol.*, vol. 10, no. 3, pp. 1578–1588, Jul. 2022, doi: 10.29130/dubited.1058467.