Journal of Materials and Electronic Devices 2 (2025) 49-57

JOURNAL OF MATERIALS AND ELECTRONIC DEVICES

JMED

# From Defect Detection to Defect Type Recognition: A Vision Transformer-Based Hybrid Framework for Magnetic Tile Surface Defect Inspection

Vahtettin Cem BAYDOGAN

Faculty of Technology, Department of Software Engineering, Firat University, Elazig, TURKEY

Defects such as cracks, blowhole, fray, uneven, breaks, and surface irregularities that occur during the production process of magnetic tiles directly and negatively affect the performance and lifespan of electric motors. Therefore, early and accurate inspection of magnetic tiles on the production line is critically important for industrial quality control processes. However, traditional visual inspection methods have significant limitations such as high cost, susceptibility to human error, and low consistency. In this study, a two-stage deep learning framework based on Vision Transformer (ViT-Base Patch16, 224×224) is proposed for the automatic detection of magnetic tile surface defects and detailed classification of defect types. In the first stage of the proposed system, magnetic tile images were subjected to binary classification as "defective" and "free". In this stage, performance evaluation was carried out using different machine learning (ML) classifiers after ViT-based feature extraction. Experimental results showed that the highest performance was obtained by the ViT+ Multilayer Perceptron (MLP) hybrid model with 97.77% accuracy, 97.30% $F_1$-score, and 97.30% Area Under Curve (AUC) value. In the second stage, the goal was to recognize different defect types (cracks, blowhole, fray, uneven, and breaks) in a multi-class manner on the images identified as defective. The most successful results in this stage were again obtained with the ViT+MLP model, achieving 94.27% accuracy, 92.43% F-score, and 96.42% AUC. These findings demonstrate that the ViT-based global context learning capability provides high success in both detecting defects and discriminating defect types in magnetic tile surfaces. In conclusion, the proposed two-stage ViT-based hybrid approach is considered to offer an effective solution for fast, reliable, and low-cost automated surface inspection systems in magnetic tile production.

***Key words:*** *Vision Transformer, Magnetic Tile, Surface Defect, Classification*

*Submission Date: 20 September 2025*
*Acceptance Date: 29 November 2025*

**Corresponding author: vcbaydogan@firat.edu.tr*

## 1. Introduction

Electric motors are among the most critical components of modern industrial systems and have a very wide range of applications, including manufacturing, transportation, energy conversion, and domestic applications. The performance, efficiency, and lifespan of these motors are directly related to the quality of the magnetic components within their internal structure [1]. Magnetic tiles, which are among these components and are generally made from ferrite-based materials, play a fundamental role in creating stable magnetic fields within the motor. Any defect occurring on the surface of magnetic tiles during the production process can lead to decreased motor efficiency, increased energy losses, vibration and noise, and ultimately a shortened system lifespan. Therefore, controlling the surface integrity of magnetic tiles with reliable inspection mechanisms is vital for industrial quality control processes [2, 3].

Magnetic tile production is a high-volume and precise manufacturing process, during which various surface defects such as cracks, pores, surface irregularities, edge wear, and structural fractures can occur. These defects can occur at different stages due to raw material-related problems, molding errors, sintering conditions, or post-production processes. While some defects are visible to the naked eye, low-contrast or small-scale surface defects are often overlooked. These undetected defects can lead to significant reductions in final product performance and, in some cases, premature motor failure [4, 5]. Therefore, early and accurate detection of magnetic tile surface defects is crucial for reducing production costs, increasing product reliability, and maintaining quality standards.

Traditional magnetic tile inspection methods rely heavily on manual visual inspections performed by trained operators. However, this approach has significant limitations, including susceptibility to human error, subjective evaluation, fatigue, time loss, and scalability problems. Manual inspection processes become both costly and inefficient, especially in mass production lines [6]. Furthermore, classical image processing-based methods may be insufficient in distinguishing between low-contrast, complex textured, or very similar defect types. This clearly highlights the need for automated, fast, and reliable surface inspection systems that can operate under real industrial conditions.

In recent years, rapid advancements in computer vision and artificial intelligence (AI) have made it possible to develop automated solutions for industrial surface defect detection [7]. While traditional ML methods such as Support Vector Machines (SVM), k-nearest neighbors (kNN), Random Forest (RF), Logistic Regression (LR), and augmentation-based classifiers have achieved some success in surface defect analysis, these methods are largely dependent on manually designed features and cannot adequately represent complex surface patterns. Deep learning (DL), and especially Convolutional Neural Networks (CNN), have made significant progress in this field thanks to their ability to automatically learn hierarchical features from large and complex image datasets. However, since CNN-based approaches rely heavily on local receiver domains, they are limited in modeling global contextual relationships across the entire surface.

ViT architectures, developed to overcome these limitations, can effectively learn long-range relationships between image patches and the global context through self-attention mechanisms. This feature makes ViTs a powerful alternative for applications such as magnetic tile inspection, where surface defects are related not only to local texture distortions but also to the overall structure of the surface. However, the use of ViT-based approaches in magnetic tile surface defect inspection is still limited in the literature.

This study proposes a ViT-based hybrid framework for the automatic detection of magnetic tile surface defects and detailed recognition of defect types. The proposed approach follows a two-stage classification strategy, consistent with industrial quality control processes. In the first stage, magnetic tile images are classified as defective and non-defective, providing rapid prescreening. In the second stage, multi-class recognition of defect types is performed on samples identified as defective. For this purpose, the ViT-Base Patch16 (224×224) architecture is used as the primary feature extractor, and a hybrid decision-making structure is created using different ML classifiers. The performance of the proposed system is comprehensively evaluated using standard metrics such as accuracy, precision, recall, F-score, and AUC. The experimental results show that the ViT-based hybrid framework achieves high success in both defect detection and defect type recognition stages. This approach, which minimizes human intervention and offers early, reliable defect detection, is considered to provide an effective and cost-effective solution for automated surface inspection systems in magnetic tile production.

The main contributions of this study to the literature are summarized below:

- To our knowledge, a ViT-based hybrid framework has been proposed for the first time.
- A two-stage classification strategy compatible with industrial quality control processes has been developed.
- The effectiveness of global context-based feature learning in surface defect discrimination has been demonstrated using hybrid ViT+ML methods.
- A comprehensive and comparative performance analysis has been presented on a realistic magnetic tile dataset.

The remainder of this automated magnetic tile surface inspection study is organized as follows: Section 2 presents a literature review of the study. Section 3 provides detailed information about the dataset and methods used. Section 4 presents the experimental results in tabular form, followed by a comparative analysis of the results. The conclusion summarizes the findings of the study in detail.

## 2. Literature Review

Detecting magnetic tile surface defects is a critical process in industrial production because it directly affects the performance and lifespan of motors. Detecting defects on these tiles presents challenges such as texture complexity, variety of defect shapes, irregular lighting conditions, and low contrast of defect areas. Common defect types encountered in the studies reviewed include blowhole, crack, fray, break, and uneven. The literature proposes several novel approaches for magnetic tile defect detection using both conventional multiscale analysis techniques and advanced DL models.

Huang et al. presented a real-time, multimodule neural network model called MCuePush U-Net for the prominence detection of magnetic tile surface defects. This model was tested on a dataset containing blowhole and crack defect types. The architecture includes a U-Net along with a Push Network that helps generate the desired prominence maps, enabling the model to efficiently map multiple defects in low-contrast images. AUC, MAE, and Fβ-Measure were used as evaluation metrics. MCuePush U-Net achieved high results on the blowhole & crack dataset, such as 0.985 in AUC, 0.002 in MAE, and 0.795 in Fβ-Measure. Its most significant finding was that the model met the demand for real-time inspection by reducing the processing time from 0.5 seconds to 0.07 seconds per image [8].

YOLO series algorithms have been frequently developed to identify magnetic tile surface defects for object detection tasks. Ma et al. proposed the YOLOv8-AHF algorithm, based on the YOLOv8 model, to improve the detection and positioning accuracy of small objects. This algorithm uses ADConv (a combination of Atrous Convolution and Depthwise Separable Convolution) to capture general and local features in the Backbone section and HAM (Hybrid Attention Module) to enhance focusing on regions of interest in the Neck section. The positioning loss function was defined as Focal-EIoU, optimized to solve the sample imbalance problem. YOLOv8-AHF increased the mAP@0.5 value to 0.962 and the mAP@0.5:0.95 value to 0.682 on the Magnetic Tile Surface Defect Dataset. This represents a 5.9% improvement in mAP@0.5 compared to the original YOLOv8 [9]. A similar study, the YOLO-RDM algorithm proposed by Niu et al., aimed to improve the discrimination ability and learning efficiency of the YOLOv8n model. YOLO-RDM replaced the traditional convolution in the Neck network with DOConv (Depthwise Over-parameterized Convolutional Layers) and the C2f module with RPA Block (Residual Parallel Attention Mechanism). MogaNet, which collects contextual information, was used as the backbone network. This model was tested on the Magnetic Tile Surface Defect Dataset and the NEU metal surface defect dataset, reporting a mAP@0.5 value of 95.0% and demonstrating high recognition and generalization ability by keeping the inference time less than 5.6 ms [10].

In the field of unsupervised learning, Tang et al. proposed a binary Struct and Detail Reverse Distillation framework called BSDRD for unsupervised anomaly detection and localization of surface defects. This framework includes a Struct Student focusing on general structural anomalies and a Detail Student using wavelet transform to capture complex textures and subtle defects. The model also utilizes MD-GFLoss (Multidimensional Feature-Gated Fusion Loss) to improve the model's selectivity for key features and sensitivity to anomalous areas. Testing for defects such as Blowhole, Break, Crack, Fray, and Uneven on a magnetic tile dataset, BSDRD demonstrated superior performance in pixel and sample-level anomaly detection, achieving scores of 0.974 in Image AUROC, 0.802 in Pixel AUROC, and 0.793 in Pixel AUPRO [11].

In addition to deep learning-based approaches, multiscale geometric analysis (MGA) techniques and diffusion methods have also been used. Xie et al. developed a Discrete Shearlet Transform (DST)-based algorithm, an MGA method, for detecting defects in magnetic tile images. DST offers higher directional accuracy compared to other MGA methods. The method uses the Threshold Surface Method to change the coefficients, and the overall accuracy rate reached 98%, demonstrating that the Shearlet Transform makes good use of geometric properties [12].

Yang et al. proposed a Stationary Wavelet Transform (SWT) based approach for low-contrast defect detection. The method applied the Sobel process to eliminate the unbalanced background, followed by index low-pass filtering and nonlinear enhancement for noise reduction and target enhancement. In experiments conducted on 85 normal and 69 defective samples, the method achieved an accuracy rate of 92.86% with an average processing time (AOT) of 0.5190 seconds [13].

Focusing on the detection of crack defects, Li et al. used a hybrid method of Fast Discrete Curvature Transform (FDCT) and Texture Analysis (GLCM). This method was able to eliminate the effect of grinding textures and accurately extract cracks larger than 0.8 mm. Tested on 59 defective and 56 qualified samples, the accuracy rate of the method was determined to be 93.9% [14].

Finally, Ben Gharsallah and Ben Braiek proposed the Improved Anisotropic Diffusion Method for defect detection/segmentation. Compared to traditional models, the proposed model combines gradient magnitude and a novel local difference image feature to eliminate bright shapes and unwanted artifacts while homogenizing the background and sharpening defect boundaries. This method achieved a detection success rate of 95% for blowholes and 89% for cracks, and claimed to be fast enough to meet real-time inspection requirements with an average computation time (ACT) of 0.83 seconds [15].

A review of the literature reveals that research on the detection and classification of magnetic tile surface defects largely focuses on traditional image processing methods, transformation-based approaches, and CNN-based DL models. While these approaches yield successful results for certain defect types and under controlled image conditions, they mostly perform feature learning based on local receiver fields and are limited in discriminating against low-contrast, fine, and irregular defects associated with the overall surface structure. Furthermore, a significant portion of existing studies either focus solely on defect detection or offer evaluations on a limited number of defect types. In contrast, the use of ViT-based architectures, which can model long-range relationships between image patches and global contextual information, in magnetic tile surface defect control is quite limited in the literature. In particular, the insufficient investigation of two-stage and hybrid ViT frameworks that combine defect presence detection and defect type recognition indicates that the potential impact of the global representation learning advantages offered by transformer architectures in this field has not yet been comprehensively explored. This situation highlights a significant research gap in the literature regarding the evaluation of ViT-based methods for both the detection and detailed classification of magnetic tile surface defects.

### 3. Materials and Methods

In this section, information about the magnetic tile surface dataset and methods applied in the study is presented.

### 3.1. Materials

The study utilized an open-source magnetic tile surface dataset obtained from the Kaggle platform [16]. Created to reduce the burden of manual quality control in magnetic tile factories, this dataset is specifically structured to reflect industrial challenges such as texture complexity, variety in defect shapes, and varying lighting conditions. Comprising a total of 2688 images, the set includes critical defect types such as blowhole, crack, break, uneven, and fray, each defect having precise pixel-level labels. Magnetic tiles are marked "free" to indicate that they are defect-free. To improve training efficiency and offset the manual labeling workload, raw data was generated into a large data pool using the Moving Least Square deformation technique and standard enhancement methods. The data distribution in each class is shown in Fig. 3.1. In general, it was evaluated that the data set did not have a balanced distribution. The number of images in each class was also specified. In this dataset, 1904 images were defect-free, while 784 images contained 5 different types of defects. For these defects, the minimum number of frays was 64, and the maximum number of browholes was 230. In general, the distribution of defective and free data, as well as the distribution of defective classes, is unbalanced.

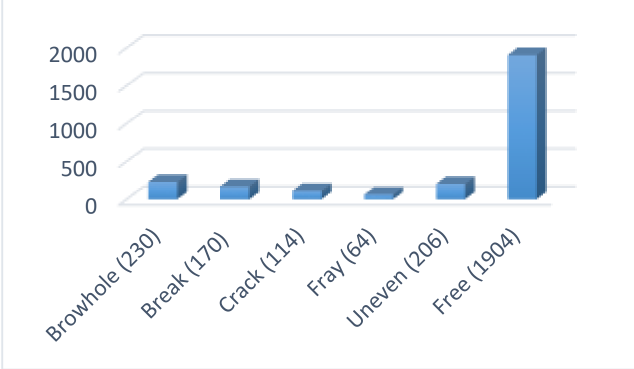



**Figure 3.1.** Classes Distribution in Magnetic Tile Surface Dataset

### 3.2. Methods

This study proposes a ViT-based hybrid approach for the detection and classification of defect types on magnetic tile surfaces. The ViT-Base-Patch16-224 architecture was used to extract distinctive features from images. The aim was to more effectively represent complex and low-contrast defect patterns on magnetic tile surfaces thanks to the ViT architecture's ability to model global contextual relationships between image patches. The deep feature vectors obtained during the feature extraction phase were fed as input to a total of 12 different ML-based classifiers for evaluation in the classification process. A comprehensive performance analysis was conducted using kNN, SVM, XGBoost, Decision Trees (DT), LR, AdaBoost, Naive Bayes (NB), Light Gradient Boosting Machine (LGBM), RF, CatBoost, ExtraTrees, and Multilayer Perceptron (MLP) algorithms. The general workflow and methodological structure of the proposed ViT-based hybrid system are presented in Fig. 3.2. This structure allows for a comparative analysis of the performance of different classifiers on ViT-based deep features, aiming to determine the most suitable classification strategies for magnetic tile surface defect inspection.

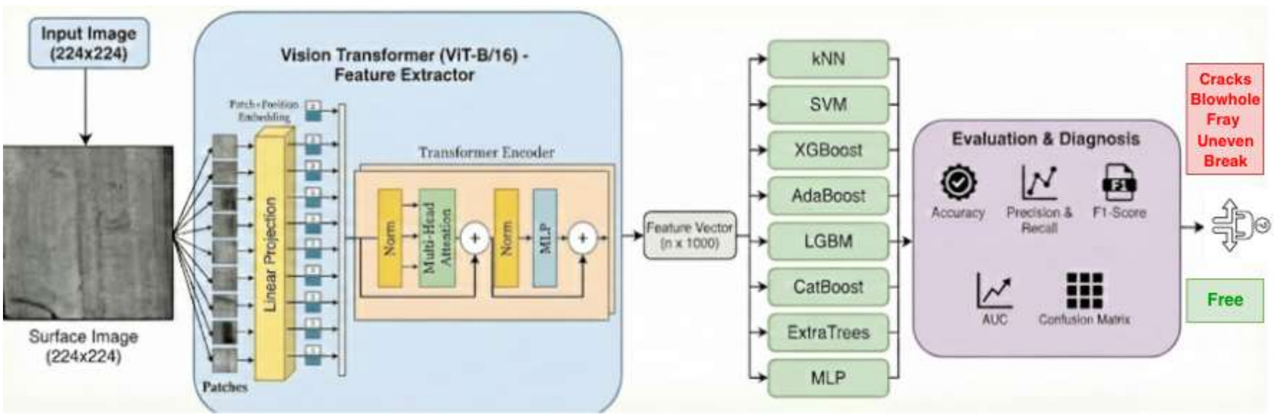



**Figure 3.2.** Pipeline of the Proposed Hybrid System for Magnetic Tile Surface Defect Inspectio

As summarized in Fig. 3.2, the proposed hybrid system begins with the processing of images representing potential defects that may occur on magnetic tile surfaces. The surface images used as input are resized to 224x224 after preprocessing steps and transferred to the ViT (ViT-Base/16) based feature extraction stage. In this stage, the images are divided into fixed-size patches, and linear projection and positional embedding operations are applied to provide input to the transformer encoder structure. Thanks to the self-attention mechanism of the ViT architecture, global contextual relationships of defect patterns on the magnetic tile surface are effectively learned, and distinctive deep feature vectors are obtained.

The deep feature vectors obtained using the ViT (ViT-Base/16) architecture in the feature extraction stage are fed as input to a total of 12 different ML-based classifiers for evaluation in the classification process. In this context, a comprehensive and comparative performance analysis was carried out using kNN, SVM, XGBoost, DT, LR, AdaBoost, NB, LGBM, RF, CatBoost, ExtraTrees, and MLP algorithms. Both binary (defective/free) and multi-class (blowhole, crack, break, uneven, and fray) defect recognition processes were performed using these classifiers. In the final stage, the performance of the proposed hybrid system was comprehensively evaluated using performance metrics such as accuracy, precision, recall, $F_1$-score, AUC, and complexity matrix. The results show that when ViT-based feature extraction is used in conjunction with state-of-the-art ML classifiers, it offers an effective and reliable solution for the detection of magnetic tile surface defects and the detailed classification of defect types.

### 3.3. Performance Evaluation Metrics

The confusion matrix is a fundamental evaluation tool used to analyze in detail the performance of AI-based classification algorithms used in the detection and classification of magnetic tile surface defects. Metrics derived from the confusion matrix, such as accuracy, precision, recall, and $F_1$-score, are widely used to statistically evaluate the performance of classification models. Considering these metrics together allows for a more comprehensive and reliable interpretation of model performance, especially in datasets where magnetic tile surface defects may have an uneven distribution. Each metric evaluates a different aspect of the model and provides important information for improving classification performance. Fig. 3.3 summarizes the general structure of the confusion matrix for the classification of magnetic tile surface defects.

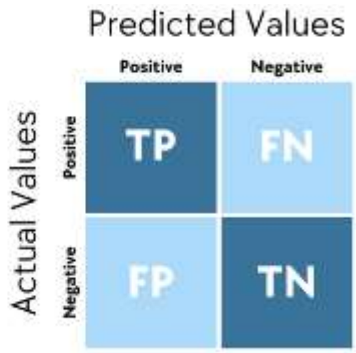



**Figure 3.3** Confusion Matrix for Model Evaluation

$$Accuracy = \left(\frac{(TP + TN)}{(TP + FP + FN + TN)}\right) \quad (1)$$

$$Precision = \left(\frac{TP}{(TP + FP)}\right) \quad (2)$$

$$Recall = \left(\frac{TP}{(TP + FN)}\right) \quad (3)$$

$$F1 - score = 2\, x \left(\frac{Precision\ x\ Recall}{Precision + \ Recall}\right) \quad (4)$$

In this context, the accuracy metric, defined by Eq. 1, expresses the overall accuracy of the model's ability to classify magnetic tile images correctly; while the precision metric, defined by Eq. 2, is a critical metric for reducing the impact of false positives, particularly in samples that need to be classified as flawless. The sensitivity metric defined in Eq. 3 evaluates the model's ability to correctly detect all defective magnetic tile samples. Finally, the F1-score given in Eq. 4 provides a balanced performance measure for surface defect detection problems where both metrics are important, taking into account both sensitivity and accuracy values.

### 4. Experimental Results and Discussion

In this experimental study, a ViT-based hybrid approach was first applied to automatically classify defective and free samples from magnetic tile surface images. Distinguishing deep features were extracted from the images using the ViT-Base-Patch16-224 architecture, and the resulting feature vectors were evaluated with 12 different ML-based classifiers. These classifiers consisted of kNN, SVM, XGBoost, DT, LR, AdaBoost, NB, LGBM, RF, CatBoost, ExtraTrees, and MLP algorithms. In the experimental evaluations, 80% of the dataset was used for training and 20% for testing, and all classifiers were tested under equal conditions. Model performance was evaluated using accuracy, precision, recall, $F_1$-score, and AUC metrics. All performance results related to magnetic tile surface defect detection are presented in Table 4.1.

**Table 4.1**. Experimental Findings for binary classification

*(Acc.: Accuracy, Pre.: Precision, Rec.: Recall, $F_1$: $F_1$-Score, Alg.: Algorithms, TX.: Transformer)*

| TX. | Alg. | Performance Evaluation Criteria | | | | |
|---|---|---|---|---|---|---|
| | | Acc. | Pre. | Rec. | $F_1$ | AUC |
| ViT-base-patch16-224 | kNN | 0.9349 | 0.9441 | 0.8979 | 0.9172 | 0.8979 |
| | SVM | 0.9145 | 0.9302 | 0.8629 | 0.8884 | 0.8629 |
| | XGBoost | 0.9517 | 0.9604 | 0.9228 | 0.9392 | 0.9228 |
| | DT | 0.8773 | 0.8509 | 0.8535 | 0.8521 | 0.8535 |
| | LR | 0.9572 | 0.9558 | 0.9399 | 0.9474 | 0.9399 |
| | AdaBoost | 0.8848 | 0.9070 | 0.8138 | 0.8444 | 0.8138 |
| | NB | 0.7007 | 0.7031 | 0.7456 | 0.6879 | 0.7456 |
| | LGBM | 0.9517 | 0.9654 | 0.9191 | 0.9386 | 0.9191 |
| | RF | 0.9312 | 0.9391 | 0.8934 | 0.9125 | 0.8934 |
| | CatBoost | 0.9312 | 0.9441 | 0.8897 | 0.9117 | 0.8897 |
| | ExtraTrees | 0.9052 | 0.9337 | 0.8413 | 0.8730 | 0.8413 |
| | MLP | **0.9777** | **0.9730** | **0.9730** | **0.9730** | **0.9730** |

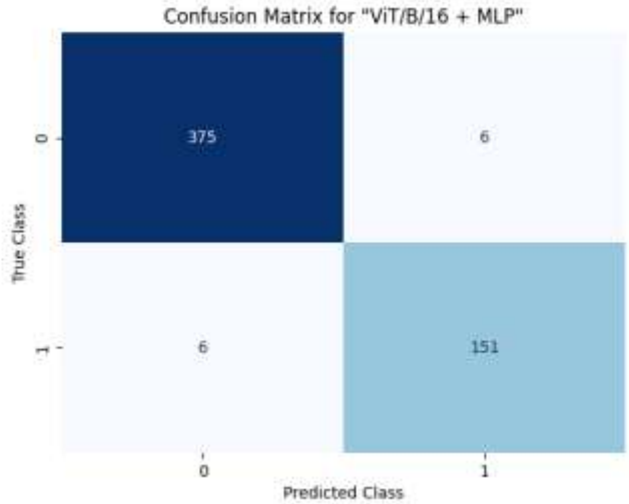
Table 4.1 shows that the performance obtained by combining ViT-based feature extraction with different ML classifiers varies significantly depending on the type of classifier. A total of 12 different experimental scenarios were evaluated in this study, and the discrimination capacity of each classifier on ViT-based features was analyzed using statistical measures. The results reveal that classification performance is strongly dependent not only on the feature extraction architecture used but also on the decision-making mechanism of the chosen classifier.

A detailed analysis of the results presented in Table 4.1 shows that the MLP classifier achieved the highest performance in the magnetic tile surface defect detection problem. The MLP model trained with ViT-based features exhibited the most balanced and highest success among all classifiers, with 97.77% accuracy, 97.30% $F_1$-score, and 0.9730 AUC. This demonstrates that MLP can effectively learn the high-dimensional and discriminatory feature space derived by ViT and successfully model complex class boundaries.

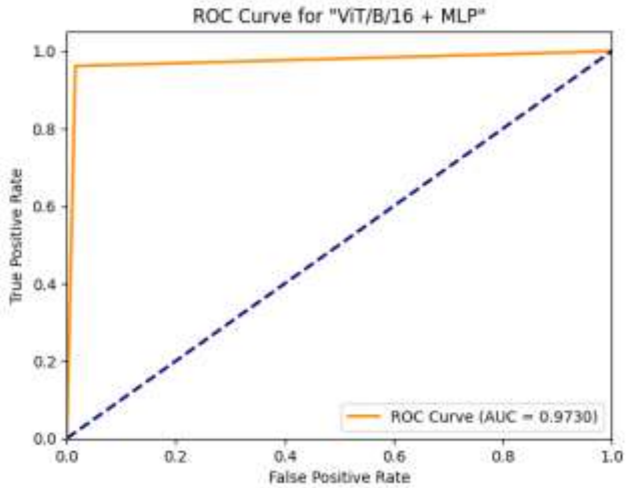
LR and XGBoost algorithms also exhibited high performance. The LR model, with 95.72% accuracy and 94.74% $F_1$-score, demonstrated that linear decision boundaries can provide strong discrimination even on ViT-based features. XGBoost demonstrated that ensemble-based methods offer an effective alternative in the magnetic tile defect detection problem, with 95.17% accuracy and a 93.92% $F_1$-score. Similarly, LGBM, RF, and CatBoost algorithms exhibited stable performance with high accuracy and balanced sensitivity values.

The kNN algorithm offered satisfactory performance with 93.49% accuracy and a 91.72% $F_1$-score, successfully capturing similarity relationships between samples in the ViT-based feature space. While the SVM model produced strong results, particularly in terms of accuracy, it tended to miss some defective samples due to a relative decrease in sensitivity. The DT, AdaBoost, and ExtraTrees models showed moderate performance, while the Naive Bayes algorithm lagged behind other methods in terms of accuracy and $F_1$-score. This suggests that the Naive Bayes' assumption of feature independence may be limiting for the complex and correlated feature structures derived by ViT.

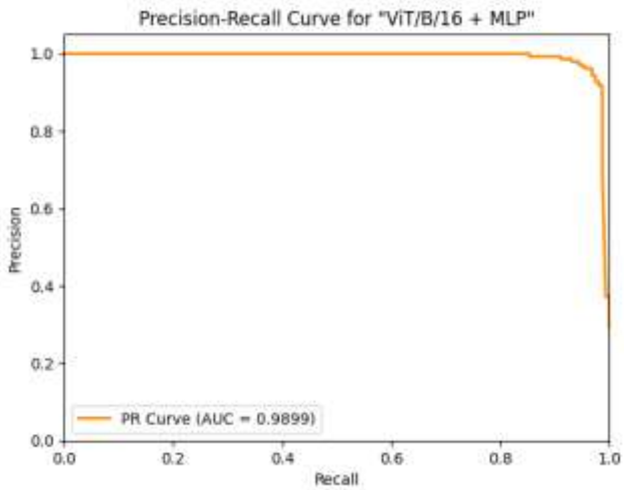
Overall, the results show that ViT-based deep feature extraction, when combined with appropriate ML classifiers, provides high accuracy, strong generalizability, and stable performance in binary classification of magnetic tile surface defects. Methods such as MLP, XGBoost, and LR, in particular, stand out in terms of both accuracy and AUC metrics, making them strong candidates for industrial quality control applications. Fig. 4.1 shows the confusion matrix, ROC, and Precision-Recall Curves for the highest-performing ViT/B/16+MLP hybrid approach in binary classification.




**a)** Confusion Matrix for ViT/B/16+MLP Hybrid Model




**b)** ROC Curve for ViT/B/16+MLP Hybrid Model




**c)** Precision-Recall Curve for ViT/B/16+MLP Hybrid Model

**Figure 4.1 a)** Confusion Matrix, **b)** ROC Curve, and **c)** Precision-Recall Curve for the ViT/B/16+MLP Hybrid Model in Binary Classification

The next stage aimed to automatically recognize five different defect types (e.g., blowhole, crack, break, uneven, and fray) observed on magnetic tile surfaces. In the multi-class classification scenario, distinctive deep features were extracted from the images using the ViT-Base-Patch16-224 architecture, and the resulting feature vectors were evaluated with 12 different ML-based classifiers. All experiments were conducted under equal conditions, with 80% of the dataset for training and 20% for testing, as in the binary classification phase. Model performances were evaluated using accuracy, precision, recall, $F_1$-score, and AUC metrics, and the results are presented in Table 4.2.

**Table 4.2**. Experimental Findings for multi-class classification

*(Acc.: Accuracy, Pre.: Precision, Rec.: Recall, $F_1$: $F_1$-Score, Alg.: Algorithms, TX.: Transformer)*

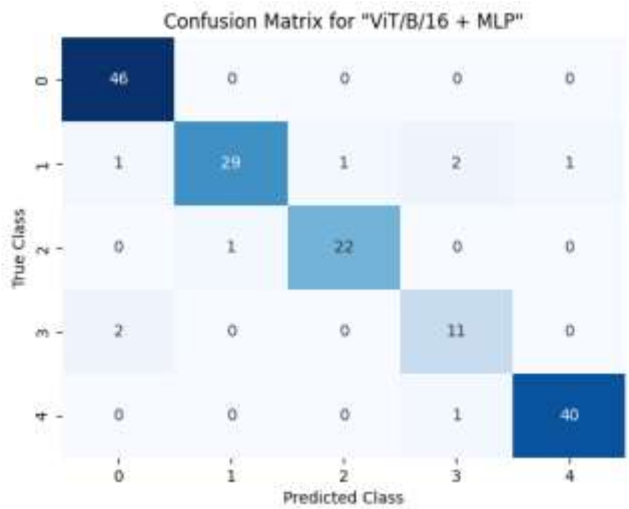
| TX. | Alg. | Performance Evaluation Criteria | | | | |
|---|---|---|---|---|---|---|
| | | Acc. | Pre. | Rec. | $F_1$ | AUC |
| ViT-base-patch16-224 | kNN | 0.8726 | 0.8716 | 0.8687 | 0.8625 | 0.9204 |
| | SVM | 0.8280 | 0.8486 | 0.8171 | 0.8298 | 0.8925 |
| | XGBoost | 0.8662 | 0.8639 | 0.8556 | 0.8593 | 0.9164 |
| | DT | 0.6497 | 0.6215 | 0.6342 | 0.6236 | 0.7811 |
| | LR | 0.9363 | 0.9139 | 0.9219 | 0.9163 | 0.9602 |
| | AdaBoost | 0.5414 | 0.3243 | 0.4126 | 0.3616 | 0.7134 |
| | NB | 0.6497 | 0.6434 | 0.6814 | 0.6522 | 0.7811 |
| | LGBM | 0.8854 | 0.8683 | 0.8653 | 0.8662 | 0.9283 |
| | RF | 0.8726 | 0.8993 | 0.8512 | 0.8707 | 0.9204 |
| | CatBoost | 0.8344 | 0.8549 | 0.8061 | 0.8255 | 0.8965 |
| | ExtraTrees | 0.8917 | 0.9004 | 0.8727 | 0.8846 | 0.9363 |
| | MLP | **0.9427** | **0.9247** | **0.9262** | **0.9243** | **0.9642** |

Table 4.2 shows that the multi-class defect type recognition problem is more challenging than the binary classification scenario, and more significant differences emerge in classifier performance. Nevertheless, it has been observed that ViT-based feature extraction, when combined with different classifiers, can achieve high success in distinguishing complex and visually similar defect types. The results reveal that the suitability of the classifier's decision mechanism to the multi-class problem structure plays a critical role in performance.

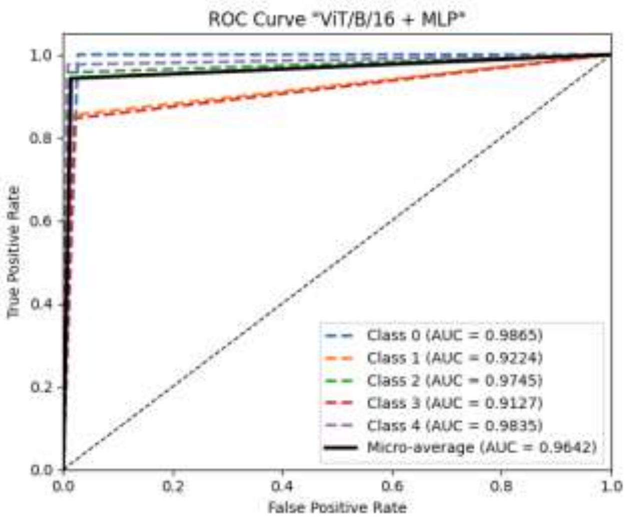
MLP classifier achieved the highest performance in the multi-class magnetic tile defect type recognition problem. The MLP model trained with ViT-based features yielded the most balanced and successful results among all classifiers, with 94.27% accuracy, a 92.43% $F_1$-score, and an AUC of 0.9642. This demonstrates that MLP can successfully model complex multi-class decision bounds by effectively learning high-dimensional and globally contextualized features derived by ViT.

The LR algorithm also exhibited remarkable performance in the multi-class classification problem. The LR model achieved the second-best result with 93.63% accuracy and a 91.63% $F_1$-score, showing that it significantly benefits from the discriminative power of ViT-based features despite having linear decision bounds. Similarly, the ExtraTrees (89.17% accuracy, 88.46% $F_1$) and LGBM (88.54% accuracy, 86.62% $F_1$) algorithms demonstrated that ensemble-based approaches offer strong and stable alternatives in the multi-class defect recognition problem.

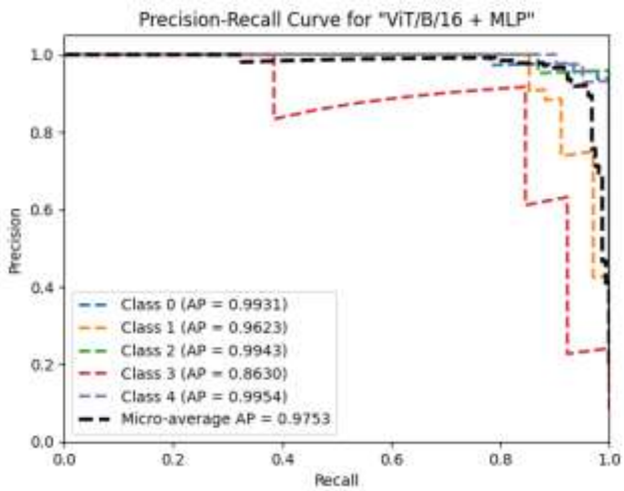
The kNN, RF, and XGBoost classifiers performed in the 86–88% accuracy range, capturing inter-class similarities to a certain extent in the ViT-based feature space. While the RF and kNN algorithms produced balanced results in terms of sensitivity and accuracy, inter-class confusion was observed due to the visual similarity of some defect types.




**a)** Confusion Matrix for ViT/B/16+MLP Hybrid Model




**b)** ROC Curve for ViT/B/16+MLP Hybrid Model




**c)** Precision-Recall Curve for ViT/B/16+MLP Hybrid Model

**Figure 4.2 a)** Confusion Matrix, **b)** ROC Curve, and **c)** Precision-Recall Curve for the ViT/B/16+MLP Hybrid Model in Multi-Class Classification

The SVM and CatBoost models offered moderate performance, while the DT, AdaBoost, and Naive Bayes algorithms showed significant performance losses in the multi-class problem structure. In particular, AdaBoost's low accuracy and $F_1$-score values indicate that its weak learner-based structure does not provide sufficient representational power for complex and unbalanced multi-class data structures. Fig. 4.2 shows the confusion matrix, ROC, and Precision-Recall Curves for the highest-performing ViT/B/16+MLP hybrid approach in multi-class classification.

Overall, the findings demonstrate that ViT-based deep feature extraction, when used in conjunction with appropriate ML classifiers, achieves high accuracy in the detailed recognition of different defect types on magnetic tile surfaces. Methods such as MLP and LR, in particular, stand out in terms of both accuracy and AUC (Accuracy, Utilization, and Reliability) metrics, and are considered powerful and reliable classifiers for multi-class magnetic tile defect type recognition problems.

## Results

This study proposes a ViT -based hybrid classification approach to address the problem of automatically detecting and recognizing surface defects in magnetic tiles, which play a critical role in industrial production processes. Defects such as blowhole, crack, break, uneven, and fray on magnetic tile surfaces directly and negatively impact the performance, energy efficiency, and lifespan of electric motors. Therefore, early and reliable detection of defects in the production line is of great importance for the sustainability of quality control processes. Traditional manual inspection and classical image processing methods may be insufficient to meet these requirements due to high costs, time loss, and susceptibility to human error.

In this study, the ViT-Base-Patch16-224 architecture was used to extract distinctive features from magnetic tile surface images, and the obtained deep features were evaluated with 12 different ML-based classifiers. Experimental studies were carried out within a two-stage structure. In the first stage, magnetic tile images were subjected to binary classification as defective and free. The results obtained showed that ViT-based feature extraction offers a very strong representation for this problem. In particular, the MLP classifier outperformed all other classifiers, achieving the highest and most stable performance with 97.77% accuracy, 97.30% $F_1$-score, and 0.9730 AUC. This result demonstrates that the ViT architecture can effectively capture global contextual information and that MLP can successfully learn this high-dimensional feature space.

In the second stage, the goal was to recognize five different defect types in a multi-class manner on magnetic tile images identified as defective. In the multi-class classification scenario, despite the more complex problem structure, the proposed hybrid approach exhibited high performance. In this stage, the MLP model achieved the most successful results with 94.27% accuracy, 92.43% $F_1$-score, and 0.9642 AUC. LR and ensemble-based methods (LGBM, ExtraTrees, RF) have also demonstrated high accuracy and balanced sensitivity values. In contrast, methods such as NB, AdaBoost, and DT have been observed to exhibit limited performance in the multi-class problem structure.

The results show that classification performance is strongly dependent not only on the feature extraction architecture used but also on the decision mechanism of the chosen classifier. ViT-based deep features, especially when combined with multilayer and ensemble-based classifiers, have been shown to provide high generalizability and stability in both the detection and detailed defect type recognition of magnetic tile surface defects. This strengthens the applicability of the proposed approach in real industrial quality control scenarios.

In conclusion, the ViT-based hybrid framework presented in this study offers an effective, fast, and low-cost solution for the automated inspection of magnetic tile surface defects. The proposed method has the potential to contribute to more reliable and consistent quality control processes in production lines by minimizing human intervention. Future studies plan to utilize explainable artificial intelligence (XAI) methods (e.g., Grad-CAM) to improve the interpretability of model decisions and to conduct tests on larger datasets encompassing diverse production conditions. Furthermore, the end-to-end classification and real-time industrial system integration of ViT-based architectures is considered a key direction for future research.

## References

**[1]** Huang, Y., Huang, Z., & Jin, T. (2024). DEU-Net: a multi-scale fusion staged network for magnetic tile defect detection. Applied Sciences, 14(11), 4724.

**[2]** Liang, W., & Sun, Y. (2022). ELCNN: A deep neural network for small object defect detection of magnetic tile. IEEE Transactions on Instrumentation and Measurement, 71, 1-10.

**[3]** Liu, T., & Ye, W. (2022). A semi-supervised learning method for surface defect classification of magnetic tiles. Machine Vision and Applications, 33(2), 35.

**[4]** Verma, J., Parsad, D., Kansal, I., Popli, R., Khullar, V., & Kumar, R. (2024). Application of Deep Neural Networks in the Detection of Surface Defects of Magnetic Tiles through the Analysis of Digital Images. In Computational Intelligence and Mathematical Applications (pp. 6-10). CRC Press.

**[5]** Zhong, Z., Wang, H., & Xiang, D. (2022). Small defect detection based on local structure similarity for magnetic tile surface. Electronics, 12(1), 185.

**[6]** Li, Y., & Fang, J. (2023). Detection of surface defects of magnetic tiles based on improved YOLOv5. Journal of Sensors, 2023(1), 2466107.

**[7]** Xie, L., Xiang, X., Xu, H., Wang, L., Lin, L., & Yin, G. (2020). FFCNN: A deep neural network for surface defect detection of magnetic tile. IEEE Transactions on Industrial Electronics, 68(4), 3506–3516.

**[8]** Huang, Y., Qiu, C., Guo, Y., Wang, X., & Yuan, K. (2018). Surface defect saliency of magnetic tile. IEEE International Conference on Automation Science and Engineering (CASE), 612–617.

**[9]** Ma, C., Pan, Y., & Chen, J. (2025). Surface defect detection of magnetic tiles based on YOLOv8-AHF. Electronics, 14(14), 2857.

**[10]** Niu, W., Lv, C., Zhang, E., & Wei, Z. (2025). YOLO-RDM: A high accuracy and efficient algorithm for magnetic tile surface defect detection with practical applications. Plos one, 20(7), e0328815.

**[11]** Tang, J., Zhang, A., & Liu, W. (2025). A novel dual-student reverse knowledge distillation method for magnetic tile defect detection. Scientific Reports, 15, 27279.

**[12]** Xie, L., Lin, L., Yin, M., Meng, L., & Yin, G. (2016). A novel surface defect inspection algorithm for magnetic tile. Applied Surface Science, 375, 118–126.

**[13]** Yang, C., Liu, P., Yin, G., Jiang, H., & Li, X. (2016). Defect detection in magnetic tile images based on stationary wavelet transform. NDT & E International, 83, 47–56.

**[14]** Li, X., Jiang, H., & Yin, G. (2014). Detection of surface crack defects on ferrite magnetic tile. NDT & E International, 62, 6–13.

**[15]** Ben Gharsallah, M., & Ben Braiek, E. (2021). Defect identification in magnetic tile images using an improved nonlinear diffusion method. Transactions of the Institute of Measurement and Control, 43(6), 1–12.

**[16]** Dataset: Magnetic Tile Surface Defects, Kaggle. (2022). https://www.kaggle.com/datasets/alex000kim/magnetic-tile-surface-defects