

PI-TransUNet: A Physics-Informed TransUNet Model for Steel Plate Defect Detection in Electromagnetic Tomography

Xianglong Liu

15111018@bjtu.edu.cn

Zhengzhou University of Light Industry

Kun Zhang

Zhengzhou University of Light Industry

Ying Wang

Zhengzhou University of Technology

Huilin Feng

Zhengzhou University of Light Industry

Nan Wang

Zhengzhou University of Light Industry

Research Article

Keywords: Electromagnetic tomography (EMT), Defect detection, Physics-informed neural network, TransUNet

Posted Date: August 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9934626/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Journal of Nondestructive Evaluation on October 11th, 2026. See the published version at <https://doi.org/10.1007/s10921-026-01446-3>.

PI-TransUNet: A Physics-Informed TransUNet Model for Steel Plate Defect Detection in Electromagnetic Tomography

Xianglong Liu^{a,*}, Kun Zhang^a, Ying Wang^b, Huilin Feng^a, Nan Wang^a

^aSchool of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou, 450002, China

^bSchool of Mechatronics and Vehicle Engineering, Zhengzhou University of Technology, Zhengzhou, 450044, China

*Corresponding author email: 15111018@bjtu.edu.cn

Abstract

Steel plate defect inspection based on electromagnetic tomography (EMT) has attracted increasing attention because of its ability to perform non-invasive imaging without direct contact with the inspected object. Nevertheless, accurate image reconstruction remains difficult owing to the highly ill-conditioned inverse mapping and the diffuse nature of electromagnetic fields, which often result in distorted defect boundaries and unstable reconstructions. Although deep neural networks have demonstrated considerable potential in EMT imaging, their predictions may deviate from underlying physical principles when trained solely from data. To improve both reconstruction accuracy and physical reliability, a physics-guided image reconstruction framework named PI-TransUNet is developed. The framework integrates a TransUNet-based inverse solver with a pretrained differentiable forward surrogate model, forming a closed-loop learning architecture. Spatial positional information is incorporated into the measurement sequence to strengthen topological awareness, while a physics-consistency objective is introduced to regularize the reconstruction process. Furthermore, a progressive weighting strategy is employed to stabilize optimization and facilitate the incorporation of physical constraints. Numerical simulations and laboratory experiments conducted on a nine-coil planar EMT platform demonstrate that the proposed approach achieves superior defect localization, enhanced boundary preservation, and improved robustness against measurement noise. Quantitative results indicate consistently high structural similarity and low reconstruction error, highlighting the effectiveness of PI-TransUNet for EMT-based steel plate defect imaging.

Keywords: Electromagnetic tomography (EMT), Defect detection, Physics-informed neural network, TransUNet.

1. Introduction

Electromagnetic tomography (EMT) is an imaging technique that utilizes electromagnetic coupling measurements to infer the internal property distribution of an object. By deploying multiple transmitting and sensing coils around the inspection domain, the system acquires boundary electromagnetic signals associated with changes in material characteristics. These measurements can then be used to reconstruct conductivity or permeability distributions inside the target region. Owing to its contactless operation, absence of ionizing radiation, economical hardware requirements, and adaptability to harsh environments, EMT has been increasingly investigated for applications including defect inspection of metallic structures, multiphase flow visualization, medical diagnosis, and industrial process supervision [1-6].

From a mathematical perspective, EMT imaging is essentially a typical nonlinear inverse problem. According to Maxwell's equations, under low frequency or quasi-static conditions, the spatial distributions of conductivity and permeability within the target region govern the propagation and

coupling characteristics of the electromagnetic field generated by the excitation coils. Due to the soft-field nature of EMT, the diffusion and propagation behavior of the electromagnetic field within the medium leads to pronounced spatial non-uniformity in measurement sensitivity, which in turn introduces strong nonlinearity and severe ill-posedness in the relationship between boundary measurements and internal physical parameters. Consequently, the imaging quality is largely determined by the modeling ability and solution strategy of the reconstruction algorithm.

To improve reconstruction quality in EMT systems, numerous algorithms have been developed over the past decades. These approaches can be broadly categorized into three groups: conventional model-based reconstruction methods, deep-learning-driven imaging frameworks, and recent physics-guided learning strategies that incorporate prior electromagnetic knowledge into network optimization.

In early studies, EMT image reconstruction relied primarily on analytical methods and iterative optimization. Among the analytical approaches, Linear Back-Projection (LBP) is representative, providing a coarse estimate of the conductivity or permeability distribution through a single back-projection operation [7]. However, due to the strong nonlinearity and ill-posedness of EMT, such methods commonly suffer from severe artifacts, blurred boundaries, and limited quantitative accuracy. To overcome these limitations, researchers have proposed a variety of iterative optimization methods, including Landweber iteration [8], the Conjugate Gradient (CG) method [9] and Iterative nonlinear Tikhonov algorithm [10]. These methods gradually approach the true physical property distribution by constructing an objective function and iteratively minimizing the residual between the predicted value of the forward model and the actual measured data in the iterative process. With the development of optimization theory, sparse regularization algorithms based on L1-Lp norm have been introduced to improve image resolution [11]. A frequency-constrained sparse Bayesian learning method has been developed for multi-frequency EMT in acute stroke detection [12]. In addition, the optimization strategy based on the Proximal Gradient Method has been applied to carbon fiber cable defect detection, effectively improving modeling ability for complex structures [13]. Although these methods improve reconstruction quality to a certain extent by introducing prior constraints, they fundamentally depend on repeated iterations, and continue to suffer from high computational complexity, slow convergence speed and sensitivity to regularization parameters.

In recent years, deep learning methods have provided new avenues for solving the highly nonlinear mapping inherent in EMT. For different tomography tasks, researchers have designed a variety of deep network architectures. The V-Net has been employed to improve reconstruction accuracy in electrical tomography [14]. The SegNet has been used to model spatial structure information in three-dimensional resistivity imaging [15]. The Res-SEUnet has been proposed to suppress artifacts and enhance boundary reconstruction in electrical capacitance tomography (ECT) [16]. The sparse autoencoder combined with the radial basis function network (SSAE+RBF) and the optimized fully connected network (FC) have been used to learn the nonlinear mapping between measurements and physical property distributions for EMT [17]. The Edge-DeepLabv3+ has further improved the reconstruction quality of the defect boundary [18]. The hybrid model combining ResNet-18 and Transformer has enhanced imaging efficiency [19]. The transformer-V-Net (TV-Net) has been used for 3D EMT reconstruction [20]. The method combining CNN and Mamba2 state space model has further strengthened 3D modeling ability [21]. In addition, the models based on residual network (ResNet) have been applied to biomedical magnetic induction tomography (MIT), effectively improving reconstruction stability and accuracy [22]. Although purely data-driven methods offer significant advantages in reconstruction accuracy and computational efficiency, they are essentially implicit mapping models that depend heavily on large amounts of labeled

data and lack explicit constraints from electromagnetic physical laws, which often leads to insufficient physical consistency and limited generalization.

More recently, PINNs have emerged as a frontier direction for solving inverse problem of electrical tomography. Such methods seek to embed physical mechanisms into the learning framework, thereby compensating for the shortcomings of purely data-driven models. Jin proposed a physics-constrained deep learning reconstruction method in which forward model consistency constraints are introduced during network training, ensuring that the reconstruction simultaneously fits the data and adheres to electromagnetic physical laws, thereby effectively improving the stability and anti-noise performance of ECT imaging [23]. This dual-driven paradigm enhances both noise robustness and physical reliability and provides a promising scientific framework for solving the EMT high-precision imaging problem.

Motivated by the above considerations, this paper proposes a physics-constrained TransUNet inverse problem reconstruction framework for EMT steel plate defect imaging. The proposed method introduces a forward problem surrogate model into the inverse problem learning process. By constraining the network's predicted conductivity distribution to generate responses consistent with the observed data under the forward physical model, a deep fusion of the physical laws and data-driven learning is achieved. The main contributions of this paper are as follows:

- (1) A dual network reconstruction framework that jointly optimizes the forward and inverse problems is established.
- (2) An inverse problem network architecture that combines the fine-grained local modeling ability of U-Net with the global feature modeling ability of the Transformer is designed, effectively enhancing the representation ability of complex defects.
- (3) Physical-consistency losses and regularization constraints are introduced to realize the effective fusion of data-driven learning and the electromagnetic physical models.
- (4) Comprehensive simulation and empirical studies confirm our approach's exceptional capabilities regarding high-fidelity reconstruction, robust operational stability, and resilience to noise.

The remainder of the manuscript is arranged as follows. The electromagnetic forward model and dataset generation process are described in Section 2. Section 3 details the architecture of the proposed physics-informed reconstruction framework and its corresponding training methodology. Reconstruction results obtained from simulation and practical experiments are analyzed in Section 4. Finally, Section 5 concludes the study and discusses possible directions for future development.

2. Electromagnetic Modeling and Dataset Preparation

The theoretical framework of the planar EMT system is introduced in this section, including the governing electromagnetic equations and the properties of the corresponding reconstruction problem. The generation and preprocessing procedures of the training dataset are then presented.

2.1 Quasi-Static Electromagnetic Field Formulation

The physical model of EMT is governed by Faraday's law of electromagnetic induction and the principle of eddy current testing. Maxwell's equations comprehensively describe the interactions among electric fields, magnetic fields, and electrical currents in macroscopic electromagnetic phenomena, serving as the theoretical foundation for formulating the EMT forward problem model.

In the developed EMT platform, an excitation frequency of 200 kHz was selected for steel plate defect inspection. Since the geometric dimensions of the system are substantially smaller than the corresponding electromagnetic wavelength, the electromagnetic field exhibits pronounced quasi-static

characteristics. Under these conditions, the conduction current within the conductive medium far exceeds the displacement current, so that the displacement current term can be safely neglected and the Ampère-Maxwell law is simplified to its quasi-static form:

$$\nabla \times H \approx J = \sigma E + J_s \quad (1)$$

where H is the magnetic field intensity, J is the total current density, σ is the electrical conductivity, E is the electric field intensity, and J_s denotes the externally applied source current density.

To further simplify the solution of the three-dimensional electromagnetic field, the magnetic vector potential A is introduced and satisfies $B = \nabla \times A$. Under the assumptions that no external voltage is directly applied across the tested steel plate and that no free charge accumulates within the medium, the scalar potential gradient vanishes ($\nabla \varphi = 0$). The electric field is therefore dominated by the induced component and simplifies to:

$$E \approx -j\omega A \quad (2)$$

By combining the constitutive relations with the Coulomb gauge ($\nabla \cdot A = 0$), the governing equation of the eddy current field in an inhomogeneous medium can be derived as:

$$\nabla \times \left(\frac{1}{\mu} \nabla \times A \right) + j\omega \sigma A = J_s \quad (3)$$

This partial differential equation explicitly characterizes the nonlinear coupling relationship between the electromagnetic properties of the medium and the spatial distribution of the magnetic vector potential, and serves as the fundamental basis for the numerical solution of the forward problem.

$$U = -j\omega \oint_l A \cdot dl \quad (4)$$

Equation (4) indicates that local abrupt changes in electrical conductivity or magnetic permeability, induced by internal defects within the steel plate, perturb the spatial distribution of the magnetic vector potential, which in turn alters the amplitude and phase of the induced voltages in the detection coils. These boundary voltage measurements provide the fundamental input data for the subsequent inverse problem of image reconstruction.

2.2 Planar EMT Sensor Configuration

As depicted in Fig. 1, a planar EMT sensing platform consisting of nine circular coils arranged in a 3×3 matrix was developed for data acquisition. Each sensing coil is designed with an inner diameter of 12 mm, an outer diameter of 20 mm, and an axial height of 10 mm. The center-to-center spacing between adjacent coils is maintained at 24 mm. During operation, the system adopts a sequential excitation strategy, where each coil is activated in turn while the remaining coils serve as receivers. Consequently, nine excitation patterns are generated, producing 81 mutual-inductance measurements in a complete scanning cycle.

A steel plate of dimensions 70 mm × 70 mm × 10 mm serves as the test object. Circular full-depth defects with radii of 3 mm and 6 mm are embedded to simulate different damage conditions. When the excitation coil is energized, the generated magnetic flux induces circulating currents within the specimen. Any discontinuity introduced by a defect modifies the current pathways and consequently affects the electromagnetic responses measured by the receiver array [18]. In the numerical model, a tetrahedral finite-element mesh comprising 26,844 elements is adopted, and the average element size is maintained at 0.6 mm to ensure computational stability and solution accuracy, as shown in Fig. 2.

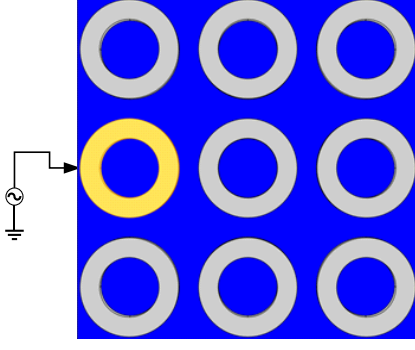


Fig. 1. Planar EMT sensor configuration.

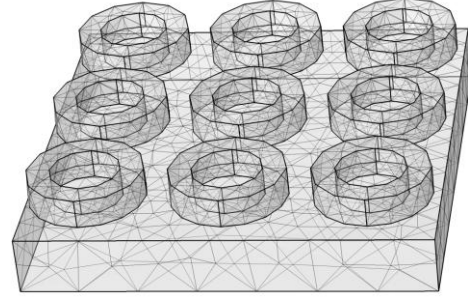


Fig. 2. Mesh generation based on finite element.

2.3 Dataset Construction and preprocessing

To acquire sufficient and physically consistent training samples, an automated data generation pipeline was developed utilizing a co-simulation framework combining COMSOL Multiphysics and MATLAB. During data acquisition, a “single-excitation, full-array reception” measurement strategy was employed. By sequentially injecting a sinusoidal excitation current with a frequency of 200 kHz and an amplitude of 10 A into each of the 9 coils, producing a measurement voltage vector V comprising 81 elements per scanning cycle:

$$V = [V_1, V_2, \dots, V_{81}]^T \quad (5)$$

To ensure model generalization and class balance, a comprehensive dataset consisting of 12,960 samples was constructed, evenly distributed across four working conditions corresponding to single, double, triple, and quadruple defects. Defects are modeled as circular through-holes with randomly assigned spatial positions. For ground truth generation, the region of interest is uniformly discretized into 324 grid elements and a pixel-level binary label vector is extracted by determining whether the geometric center of each element lies within a defect boundary (1 for defect, 0 for background). Representative defect distributions are shown in Fig.3.

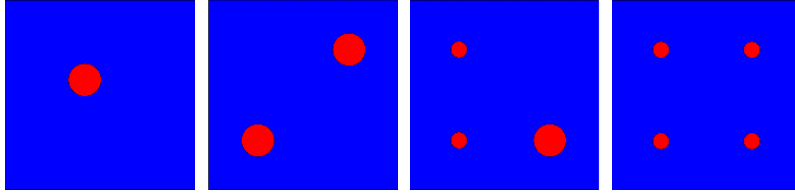


Fig.3 Representative samples of defect distributions

In the preprocessing stage, the large self-induced voltages of the excitation coils are observed to overwhelm the weaker mutual-inductance signals that carry the defect characteristics. To address this, all self-excitation channels in the measurement vector are explicitly set to zero. Min-Max normalization is then applied to eliminate dimensional discrepancies and accelerate network convergence, scaling the data into the $[0, 1]$ interval:

$$v_{norm} = \frac{v - v_{min}}{v_{max} - v_{min}} \quad (6)$$

3. Physics-Informed PI-TransUNet Architecture

This section proposes a physics-informed hybrid imaging network, termed PI-TransUNet. In terms

of an overall topological perspective, the framework adopts a generator-evaluator dual-network coupled architecture. The generator is responsible for mapping one-dimensional boundary voltage signals to high-dimensional two-dimensional conductivity images, while the evaluator acts as a physics-constraint module that verifies the physical consistency and plausibility of the generated results. The core innovations of the architecture lie in three aspects: a positional-encoding mechanism at the input layer, a global-local hybrid modeling strategy at the feature-extraction layer, and a physics-driven closed-loop mechanism at the loss optimization layer. The overall network structure is illustrated in Fig. 4.

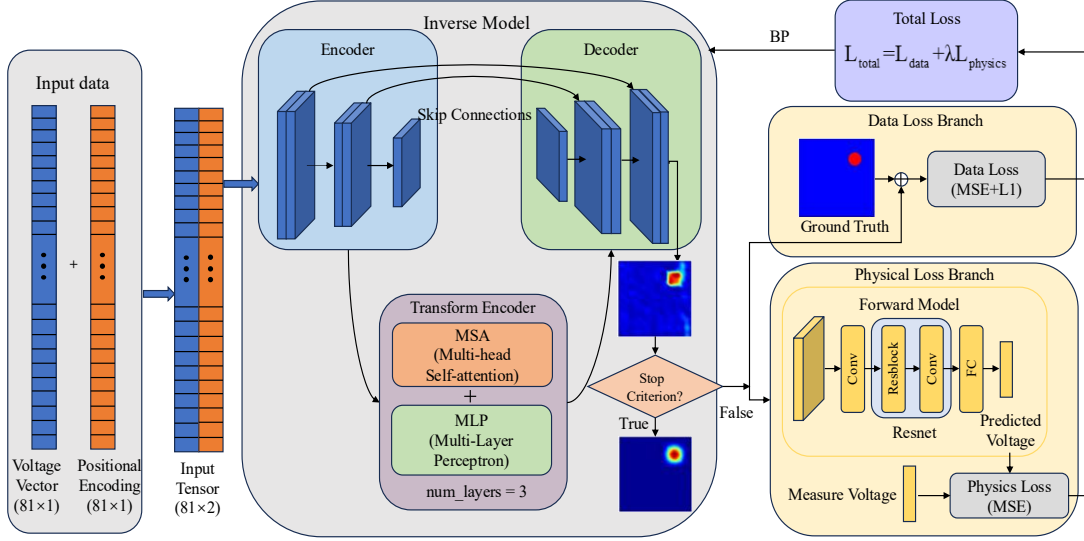


Fig. 4 PI-TransUNet network framework

3.1 Positional Encoding Strategy

In the EMT system, the 81-dimensional boundary voltage data acquired by the sensor array inherently encodes strict physical and spatial topological relationships. Conventional convolutional neural networks typically treat these data as unordered or simple one-dimensional sequences, completely neglecting the geometric topology of the sensors in physical space. However, the electromagnetic field distribution exhibits pronounced spatial structural characteristics, and the boundary voltage values are strongly correlated with their corresponding spatial locations. To enhance the network's perception of the sensors spatial layout and to assist the model in better capturing the geometric mapping between boundary measurements and internal field distributions, a positional encoding mechanism is incorporated.

Specifically, a normalized positional index vector P with the same dimension as the voltage measurement vector is constructed, in which the i -th element of this vector represents the normalized coordinate index of the i -th measurement channel within the sensor array. The normalized voltage data V and the positional encoding P are then concatenated along the channel dimension to form an input tensor X with dimensions of 81×2 :

$$X_{in} = \text{Concat}(V, P) \quad (7)$$

By introducing this mechanism, the first layer of the network simultaneously acquires two complementary streams of information: measurement amplitude and spatial coordinates, which is essential for accurate reconstruction.

3.2 TransUNet-Based Inverse Reconstruction Generator

To address the two principal challenges of blurred local features and complex global coupling relationships caused by the soft-field effect in the EMT inverse problem, conventional CNNs exhibit strong local feature extraction but are constrained by their inherently limited receptive fields, which makes it difficult to characterize the global spatial correlations induced by electromagnetic field diffusion. To simultaneously achieve fine-grained local detail recovery and global topological structure modeling, this paper designs an improved TransUNet-based inverse reconstruction generator that integrates the architectural advantages of both CNNs and Transformers. The network adopts an encoder-Transformer bottleneck-decoder architecture and aims to realize a high-precision nonlinear mapping from one-dimensional boundary measurement voltage signals to two-dimensional internal conductivity distribution images. The detailed architectural schematic is illustrated in Fig. 5.

At the network input stage, the 81-dimensional measurement voltage vector and its corresponding positional encoding are concatenated along the channel dimension to construct an input tensor of size 81×2 , in which the first dimension represents the normalized voltage amplitude and the second dimension denotes the normalized spatial coordinate index of the corresponding measurement channel. This explicit injection of spatial positional priors, substantially enhances the network's perception of the topological arrangement of the sensor array. The input tensor is then transformed into the feature space via a two-layer fully connected (FC) projection module: it is first mapped into a 512-dimensional latent space by a linear FC layer, then projected into a 324-dimensional feature vector by a second FC layer, and finally reshaped into an 18×18 two-dimensional initial feature map.

In the encoder stage, the network utilizes three successive convolutional blocks to build a hierarchical feature extraction pathway. The initial convolutional module applies 3×3 kernels with 32 output channels to capture fine-grained local texture information while maintaining the original spatial resolution of the input feature map. Subsequent spatial downsampling is achieved through strided convolutions (stride = 2), which gradually shrink the spatial dimensions of the feature maps and simultaneously increase the channel depth to 64 and then 256. This process produces rich, high-level semantic representations. During the downsampling procedure, two skip connection paths are preserved: one originates from the shallow convolutional layer and carries high-resolution edge and geometric details, while the other comes from the intermediate convolutional layer and encodes medium-scale structural information. These multi-scale features are later fused with the upsampled outputs in the decoder stage, effectively compensating for the loss of spatial information caused by downsampling operations.

At the deepest stage of the encoder, the extracted high-level feature maps are reshaped into a one-dimensional sequence and subsequently input into the Transformer bottleneck. This bottleneck is formed by three successive Transformer encoder layers, each incorporating a Multi-Head Self-Attention (MSA) module and a Feed-Forward Network (FFN). By employing the MSA mechanism, the model dynamically determines the attention weights between any pair of spatial positions across the full imaging domain:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (8)$$

where Q , K and V denote the query, key and value matrices, respectively, and d_k represents the feature dimension scaling factor. By performing attention operations in parallel across multiple heads, the network can learn complementary electromagnetic features from different representation subspaces, thereby improving its capability to model complex conductivity distributions.

Unlike conventional convolution operations that focus solely on local receptive fields, the proposed

attention-guided Transformer module establishes direct global dependencies between distant spatial positions. This mechanism fundamentally overcomes the local receptive field limitation of CNNs and enables the network to capture long-range electromagnetic coupling relationships induced by subtle conductivity perturbations. Moreover, the attention weights adaptively emphasize highly sensitive regions associated with conductivity anomalies while suppressing redundant background responses and measurement noise. As a result, the network achieves stronger robustness and reconstruction stability for scenarios involving central defects, weak-sensitivity blind zones, and complex multi-defect distributions.

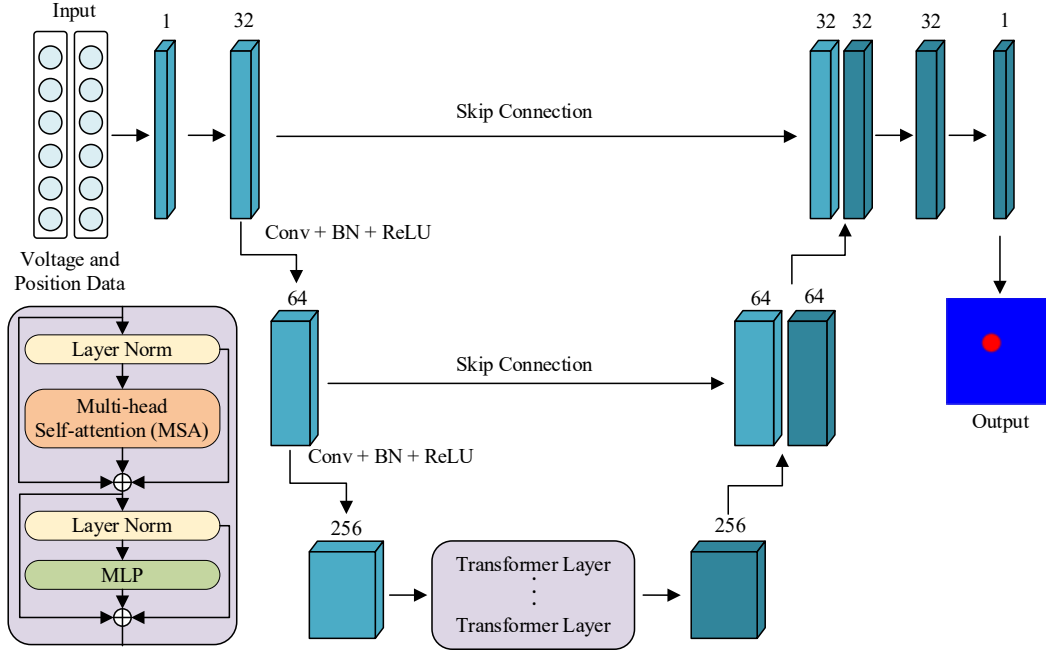


Fig. 5 TransUNet-Based Inverse Reconstruction Generator

To improve the representational capability of the Transformer features, a channel attention module is employed after the Transformer encoder. The feature maps are first compressed along the spatial dimension through global average pooling to generate channel descriptors. These descriptors are subsequently processed by a lightweight gating network consisting of two fully connected layers and a sigmoid function, producing channel-wise importance coefficients. By adaptively reweighting the feature channels, the network can highlight conductivity-related information while reducing the influence of redundant responses. As a result, more discriminative features are obtained, leading to improved reconstruction accuracy and clearer defect boundaries.

The decoding stage adopts a progressive upsampling strategy to gradually restore the spatial resolution of the feature maps. Transposed convolutions are initially employed to enlarge the deep features produced by the Transformer bottleneck. These upsampled features are then concatenated along the channel dimension with the corresponding medium-scale skip connections from the encoder, achieving the first stage of multi-scale feature fusion to reconstruct the overall structural contours of the defects. A second transposed convolution further increases the resolution and merges the result with the shallow high-resolution skip features, thereby accurately recovering fine-grained edge details. To ensure the final output precisely matches the target 18×18 grid resolution, bilinear interpolation combined with a standard convolution is applied for final feature refinement. Finally, a 1×1 convolutional layer reduces the feature channels to one, and a sigmoid activation function constrains the output values within the $[0, 1]$ range, yielding the normalized conductivity distribution image.

3.3 Physics-Guided Forward Surrogate Evaluator

Embedding a non-differentiable and computationally expensive finite element method (FEM) solver directly within the neural network is impractical. To overcome this limitation, a differentiable forward surrogate model based on an improved ResNet is proposed. This evaluator maps the two-dimensional conductivity image generated by the inverse network back into the one-dimensional boundary voltage space. It employs convolutional layers with dropout for spatial feature extraction, followed by identity mapping residual blocks that alleviate the vanishing gradient problem and preserve subtle perturbation features. After further downsampling and flattening, a fully connected layer outputs the predicted voltage vector. The architecture of the forward physical surrogate model is shown in Fig 6.

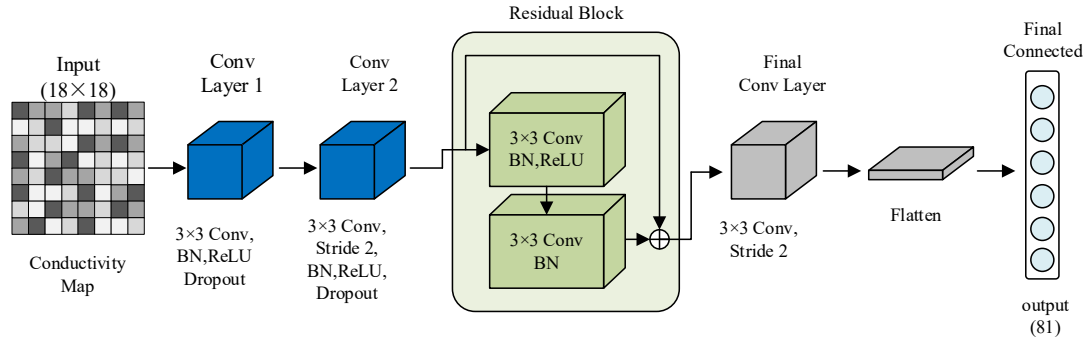


Fig. 6 Physics-Guided Forward Surrogate Evaluator

During joint training, the parameters of the evaluator (pre-trained on FEM data) are strictly frozen. It constructs a physical constraint loss by calculating the mean squared error (MSE) between the predicted and actual voltages, thereby guiding the inverse generator toward physically plausible solutions. To prevent misleading gradients caused by the highly disordered images produced in the early stages of generator training, a dynamic warm-up strategy is further introduced, in which the weight of the physical loss increases linearly from 0 to 0.5 over the first 40 epochs. This strategy substantially enhances the model's generalization and robustness in highly noisy environments.

3.4 Loss function and training strategy

By constructing a hybrid loss function that balances data mapping fidelity with electromagnetic physical consistency, this section achieves effective regularization of the severely ill-posed EMT inverse problem.

(1) Data-Driven Loss: To jointly enforce global structural consistency and local edge sharpness in the reconstructed images, a weighted combination of MSE and the L1 norm is adopted as the data-driven loss:

$$L_{data} = 0.5 \times \text{MSE}(\hat{G}, G) + 0.5 \times \text{L1}(\hat{G}, G) \quad (9)$$

(2) Physical Consistency Loss: The physical constraint loss is defined as the mean squared error between the boundary voltage predicted by the forward solver and the actual measured voltage:

$$L_{physics} = \text{MSE}(f_{fwd}(\hat{G}), V) \quad (10)$$

(3) Dynamic Weight Warm-up Strategy: During the early stages of joint training, the images generated by the inverse model exhibit strong randomness. Imposing strict physical constraints at this stage would yield highly fluctuating and misleading gradients, thereby destabilizing model convergence. To overcome this, a dynamic weight coefficient $\lambda(t)$ depending on the training iteration t is

introduced, and the total loss function is formulated as:

$$L_{total} = L_{data} + \lambda \cdot L_{physics} \quad (11)$$

During the first 40 epochs of training, the weight coefficient $\lambda(t)$ increases smoothly and linearly from 0 to a predefined target threshold of 0.5. This warm-up strategy intelligently guides the network to rapidly identify the spatial locations of defects under the guidance of the data-driven loss, and then to refine the reconstruction in the middle and later training stages under strict physical closed-loop constraints, thereby substantially enhancing the numerical accuracy and physical consistency of the reconstructed conductivity distribution.

3.5 Training Convergence Analysis

To further assess the stability and effectiveness of the proposed physics-informed training approach, the convergence characteristics of both the inverse reconstruction generator and the forward physical surrogate model were examined throughout the joint training process. Fig. 7 and 8 display the training and validation loss curves of the inverse generator and the forward evaluator, respectively.

As shown in Fig. 7, the inverse reconstruction model achieves fast convergence during the initial training phase. Both the training and validation losses decline rapidly within the first 20 epochs, demonstrating that the network quickly learns the complex nonlinear relationship between boundary voltage measurements and the internal conductivity distribution. Subsequently, the loss curves stabilize and converge smoothly. Although the validation loss is slightly higher than the training loss, the gap remains relatively small and exhibits no noticeable oscillation or divergence. This behavior indicates strong generalization ability and the absence of severe overfitting. In the later training stage, the training loss decreases to approximately 0.005, while the validation loss stabilizes around 0.015. These outcomes confirm that the hybrid Transformer-CNN architecture is capable of delivering high-fidelity conductivity image reconstruction. Furthermore, the smooth convergence also verifies the effectiveness of the dynamic warm-up strategy used for the physical consistency constraint. By gradually introducing the physical loss in the early phase, unstable gradient interference is effectively suppressed, which ultimately enhances the overall convergence robustness of the PI-TransUNet framework.

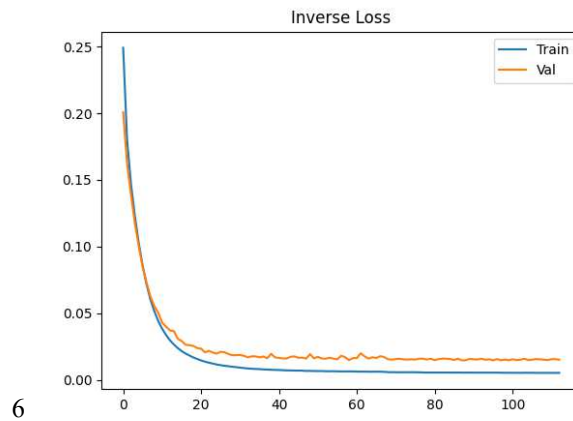


Fig. 7 Training and validation loss curves of the inverse reconstruction model.

As shown in Fig. 8, the forward physical surrogate model also exhibits strong convergence stability. The loss drops rapidly in the early training phase, indicating that the forward solver quickly learns the electromagnetic mapping from the conductivity distribution to the boundary voltage response. After around 60 epochs, both training and validation losses reach very low values. The validation loss closely

follows the training loss throughout the training process, with only minor fluctuations during the initial stage caused by random mini-batch sampling. Ultimately, the loss converges to the order of 10^{-4} , demonstrating that the proposed ResNet-based surrogate evaluator can accurately approximate the FEM-based forward solver while preserving good generalization capability. As the forward surrogate acts as the key physical constraint module in the joint optimization, its stable convergence is critical for providing reliable physical feedback to the inverse reconstruction generator. These findings confirm that the proposed physics-informed dual-network framework achieves both numerical stability and physical consistency during end-to-end training.

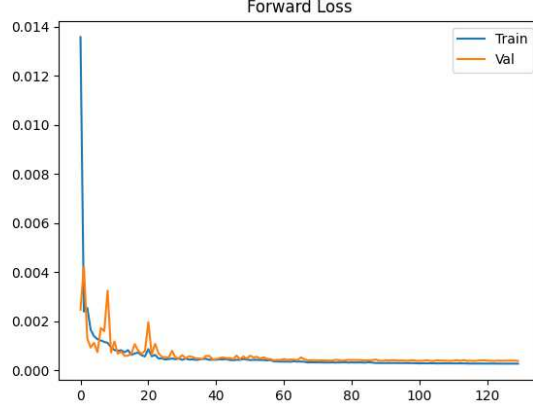


Fig. 8 Training and validation loss curves of the forward physical surrogate model.

4. Results and Discussion

4.1 Evaluation Metrics

To objectively assess reconstruction quality, both pixel-level accuracy and structural similarity are considered. Image Error (IE) is employed to quantify the normalized reconstruction error, while Correlation Coefficient (CC) is used to measure the similarity between the reconstructed and reference images. Lower IE values indicate higher reconstruction accuracy, whereas higher CC values imply better preservation of spatial structures and defect morphology. The corresponding formulations are expressed as follows:

$$IE = \frac{\|\hat{x} - x\|}{\|x\|} \quad (12)$$

$$CC = \frac{\sum_{i=1}^N (\hat{x}_i - \bar{\hat{x}})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^N (\hat{x}_i - \bar{\hat{x}})^2 \sum_{i=1}^N (x_i - \bar{x})^2}} \quad (13)$$

where N represents the total number of discretized pixels within the imaging computational domain.

The vectors x_i and $\hat{x}_i$ denote the local intensities of the i -th pixel for the ground truth and the reconstructed configurations, while $\bar{x}$ and $\bar{\hat{x}}$ designate their respective arithmetic mean values.

4.2 Ablation Study

To systematically investigate the individual contributions and synergistic effects of the proposed

modules in PI-TransUNet, an ablation study was conducted focusing on two key components: the spatial positional encoding mechanism and the physics-informed loss constraint imposed through the forward surrogate model. All experiments were performed under identical training conditions, including dataset partitioning, learning rate, batch size, and network hyperparameters. Table 1 presents a detailed breakdown of the specific configurations for the four models.

Table 1. Four model configurations

Model Configuration	Positional Encoding	Physics-informed Loss
M1	×	×
M2	√	×
M3	×	√
M4	√	√

To quantify the effectiveness of positional encoding and physics-guided regularization, four progressively enhanced model variants were designed. M1 denotes the baseline TransUNet. M2 augments the baseline with positional encoding to inject spatial priors into the measurement domain. M3 introduces a physics-consistency constraint derived from a pretrained forward model, while M4 integrates both components to establish the proposed PI-TransUNet. Reconstruction performance is evaluated using RMSE, MAE, and SSIM. RMSE and MAE characterize pixel-level reconstruction errors, whereas SSIM reflects structural similarity between reconstructed and reference images. The corresponding results are reported in Table 2.

Table 2. Quantitative results of the ablation experiments

Model Configuration	RMSE	MAE	SSIM
M1	0.1014	0.0317	0.8217
M2	0.0951	0.0336	0.8576
M3	0.0967	0.0304	0.8732
M4	0.0904	0.0276	0.9115

As shown in Table 2, both positional encoding and the physics-informed constraint individually improve the reconstruction performance. Compared with the M1 model, the M2 model reduces the RMSE from 0.1014 to 0.0951 and increases SSIM from 0.8217 to 0.8576, indicating that positional encoding enhances spatial feature representation and structural preservation. The M3 model achieves a lower MAE (0.0304) and a higher SSIM (0.8732), demonstrating that the physics-informed loss effectively improves reconstruction consistency and suppresses artifacts. The complete PI-TransUNet model achieves the best overall performance, achieving RMSE, MAE, and SSIM values of 0.0904, 0.0276, and 0.9115, respectively. Compared with M1, RMSE and MAE decrease by 10.8% and 12.9%, while SSIM increases by 10.9%.

These results indicate that the positional encoding mechanism and the physics-informed constraint provide complementary advantages. Positional encoding enhances the extraction of spatial features and long-range dependencies, whereas the physics-informed loss enforces consistency with electromagnetic forward principles. Their joint use enables PI-TransUNet to achieve improved reconstruction accuracy while preserving both structural information and physical plausibility.

4.3 Simulation results without noise

To evaluate the baseline reconstruction capability of the proposed PI-TransUNet, comprehensive simulation experiments were conducted under noise-free conditions. Fig. 9 presents the comparisons between the ground-truth conductivity distributions and the corresponding reconstructed images for

various defect configurations.

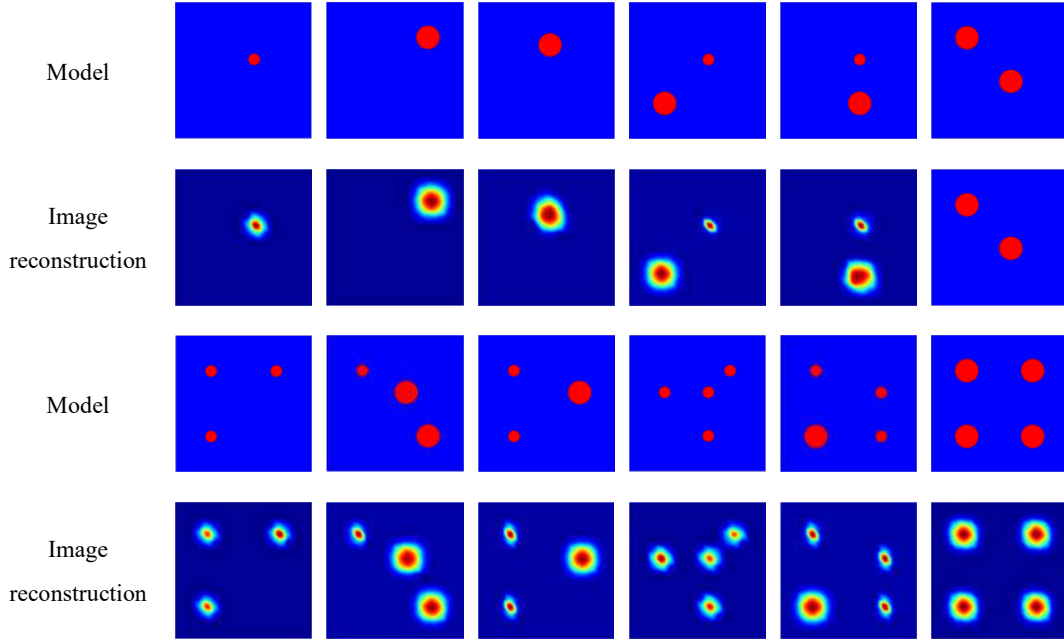


Fig. 9 Image reconstruction results under noise-free conditions.

To further validate the superiority of the proposed framework, qualitative comparisons were performed against conventional reconstruction algorithms, including LBP, Tikhonov regularization, CNN, and U-Net. Fig. 10 illustrates the reconstruction results of the different methods under multiple simulated defect distributions. As shown in Fig. 10, the LBP algorithm suffers from severe artifacts and pronounced image blurring due to the inherent ill-posedness of the EMT inverse problem. The Tikhonov regularization method partially suppresses background artifacts but still exhibits poor boundary resolution and limited structural recovery. Compared with traditional algorithms, the deep learning-based CNN and U-Net markedly improve reconstruction quality and reduce background interference. However, CNN-based reconstructions still suffer from incomplete boundary preservation and local geometric distortion, especially in multi-defect scenarios. U-Net achieves better structural continuity through its multi-scale skip connections, but still exhibits target merging and edge smoothing when reconstructing closely distributed defects. In contrast, the proposed PI-TransUNet consistently achieves the best visual reconstructions among all comparative methods. Benefiting from the Transformer-based global dependency modeling and the physics-informed joint optimization strategy, the network can accurately recover complex defect morphologies while effectively suppressing background artifacts. Even for densely distributed multi-defect structures, PI-TransUNet maintains clear edge separation and accurate geometric representation, demonstrating superior robustness and feature discrimination capability.

Table 3 quantitatively summarizes the CCs and IEs of the different reconstruction algorithms under noise-free simulation conditions. Traditional methods generally produce relatively low CC values and high IE values, indicating limited reconstruction accuracy and poor structural consistency. Among them, LBP algorithm produces the largest reconstruction errors, with IE values exceeding 0.8 in most cases, while Tikhonov regularization method provides only marginal improvements. Compared with traditional methods, the deep learning-based approaches substantially improve reconstruction performance. Both CNN and U-Net achieve higher CC and lower IE values, demonstrating enhanced feature extraction and structural recovery capabilities. Their reconstruction performance, however, still degrades in complex multi-defect scenarios. The proposed PI-TransUNet consistently achieves the best quantitative

performance across all test models. Specifically, the CC values exceed 0.89 for all defect configurations and reach a maximum of 0.948, indicating extremely high structural similarity to the ground truth, while the corresponding IE values remain below 0.22, with the minimum reduced to 0.053.

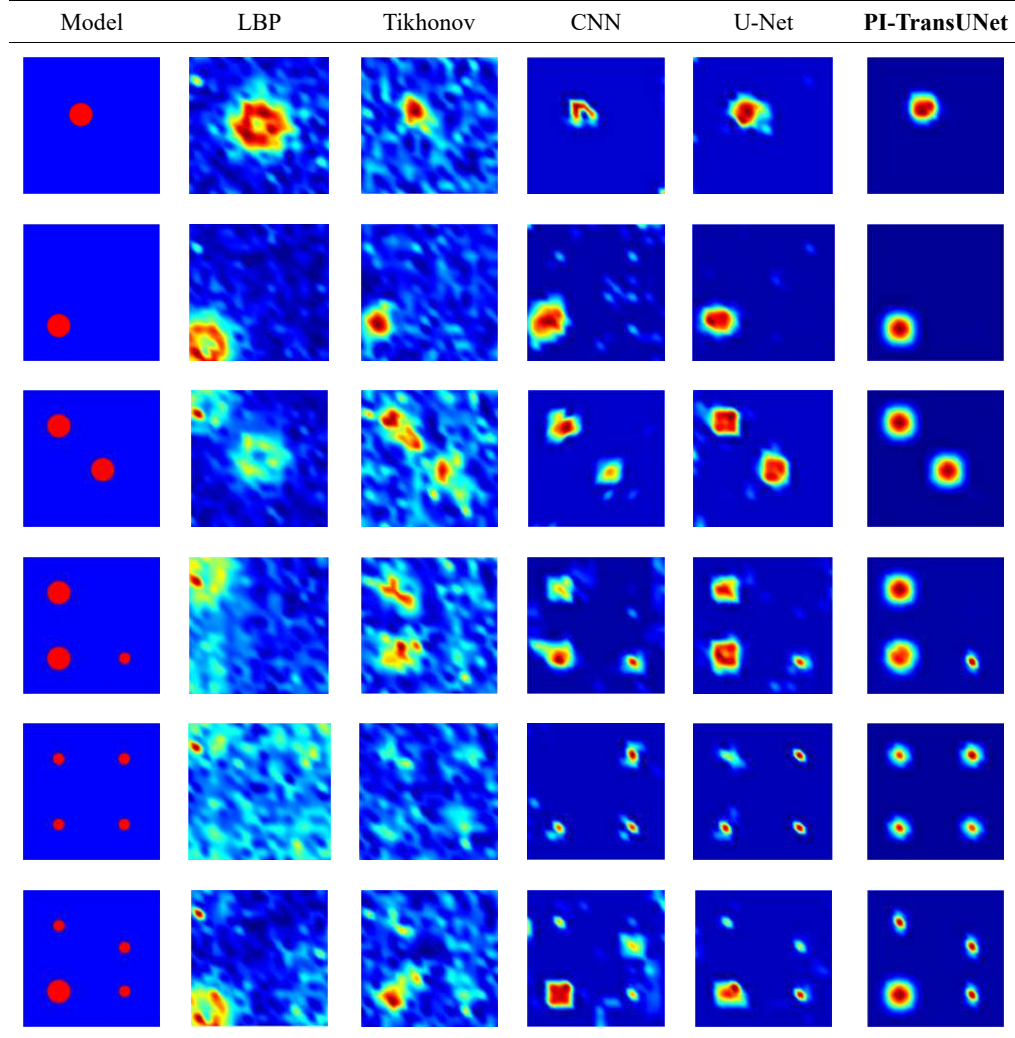


Fig. 10 Image reconstruction results of different algorithms under noise-free conditions.

Table 3. CC and IE values of noise-free simulation reconstruction

Model	Metrics	LBP	Tikhonov	CNN	U-Net	PI- TransUNet
1	CC	0.3711	0.4616	0.6752	0.7254	0.9488
	IE	0.8383	0.6489	0.5078	0.3018	0.1113
2	CC	0.4464	0.4124	0.6716	0.7819	0.8923
	IE	0.8566	0.6508	0.3857	0.2845	0.0531
3	CC	0.2893	0.4403	0.6042	0.8127	0.9425
	IE	0.8121	0.6697	0.5201	0.3463	0.1254
4	CC	0.1832	0.4504	0.6356	0.7895	0.9193
	IE	0.8585	0.6459	0.5157	0.3754	0.1563
5	CC	0.1139	0.3619	0.6175	0.7347	0.8935
	IE	0.8827	0.7017	0.5271	0.2987	0.1671
6	CC	0.0185	0.2065	0.6134	0.7547	0.9153
	IE	0.9457	0.7525	0.5581	0.3464	0.2132

These results confirm that the proposed PI-TransUNet effectively balances reconstruction accuracy and structural fidelity, achieving highly accurate and robust defect imaging under noise-free simulation conditions.

4.4 Simulation results with different noise levels

To evaluate the robustness of the proposed PI-TransUNet under complex electromagnetic interference environments, Gaussian white noise with three different signal-to-noise ratios (SNRs) was superimposed on the simulated measurement voltages. The corresponding reconstruction results are presented in Fig. 11. As shown in Fig. 11, the proposed method maintains excellent reconstruction performance under different noise conditions. Under noise-free and mild-noise (30 dB) conditions, the reconstructed images exhibit high-fidelity defect localization with virtually no visible background artifacts. As the SNRs decreases to 20 dB, slight speckle noise appears in the background, yet the spatial locations and geometric shapes of the defects remain accurately preserved. Even under severe interference at 10 dB, although background artifacts intensify and defect boundaries become slightly blurred, the proposed method can still clearly distinguish all individual defect targets without obvious target loss or substantial structural distortion. In particular, for complex multi-defect distributions containing mixed-size defects, small defect features are well preserved, demonstrating strong noise robustness and stable structural reconstruction capability.

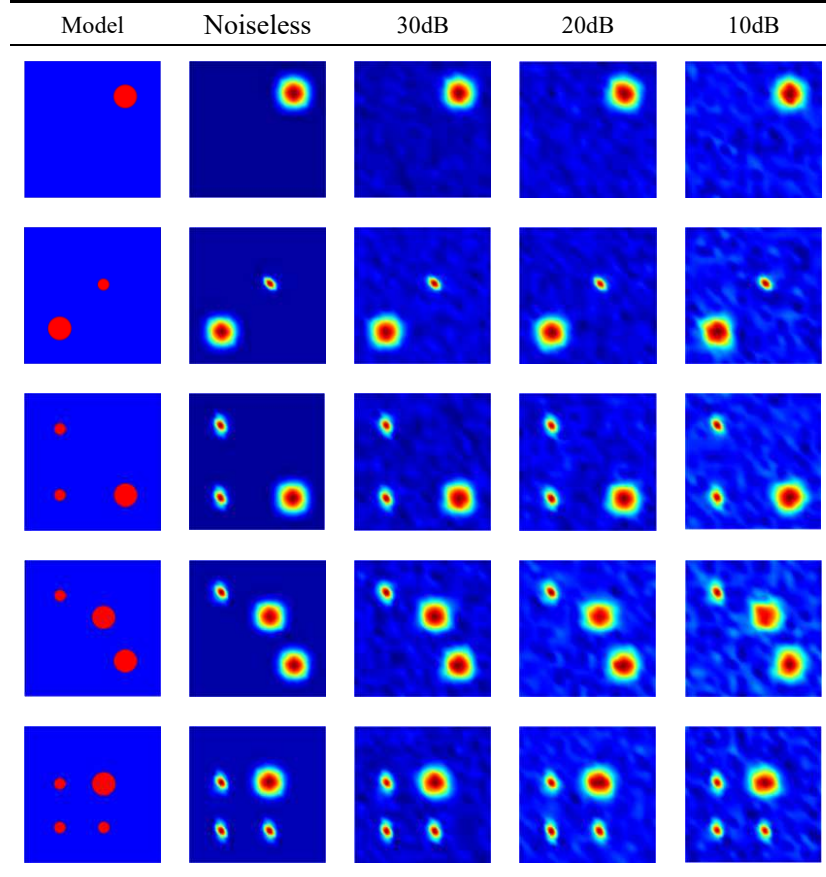


Fig. 11 Image reconstruction results under different noise levels.

The CC and IE values under different noise levels are summarized in Table 4 to quantitatively evaluate the noise robustness of the proposed method. Under noise-free conditions, the maximum CC reaches 0.9388, while the minimum IE is only 0.0971. As the noise level gradually increases, the evaluation metrics exhibit only moderate and controllable degradation, indicating strong reconstruction

stability. Even under severe 10 dB noise, the proposed method still maintains high reconstruction fidelity. For the most complex defect distribution, the CC decreases only from 0.8848 to 0.8161, while the maximum IE remains strictly below 0.21. Across all test models, the CC values consistently remain above 0.81 and the IE values remain below 0.21, confirming the superior noise robustness of the proposed physics-informed reconstruction framework.

Table 4. CC and IE values of image reconstruction under different noise levels

Model	Metrics	Noiseless	30dB	20dB	10dB
1	CC	0.9388	0.9283	0.9074	0.8848
	IE	0.0971	0.1136	0.1316	0.1482
2	CC	0.9128	0.8957	0.8618	0.8455
	IE	0.1024	0.1152	0.1477	0.1675
3	CC	0.9256	0.9143	0.9028	0.8813
	IE	0.1254	0.1383	0.1695	0.1829
4	CC	0.8921	0.8809	0.8703	0.8628
	IE	0.1392	0.1580	0.1642	0.1895
5	CC	0.8848	0.8493	0.8207	0.8161
	IE	0.1477	0.1641	0.1813	0.2077

4.5 Experimental Validation on the EMT Platform

To further validate the practical applicability of the proposed method, a custom-built nine-coil planar EMT experimental system was developed, as shown in Fig. 12. The sensor array consists of nine identical electromagnetic coils rigidly mounted on a custom-designed platform to ensure precise spatial alignment and high measurement repeatability. During operation, a Tektronix AFG1022 function generator generates a 200 kHz sinusoidal excitation signal, which is amplified through an LT1210-based power amplifier to produce a stable primary magnetic field. To minimize channel crosstalk, a dual-channel switching architecture employing two ADG406 multiplexers is adopted to provide independent excitation and reception control. The weak induced signals are further amplified by an AD8253 programmable instrumentation amplifier and acquired through a USB3200N data acquisition card for subsequent signal demodulation and image reconstruction.

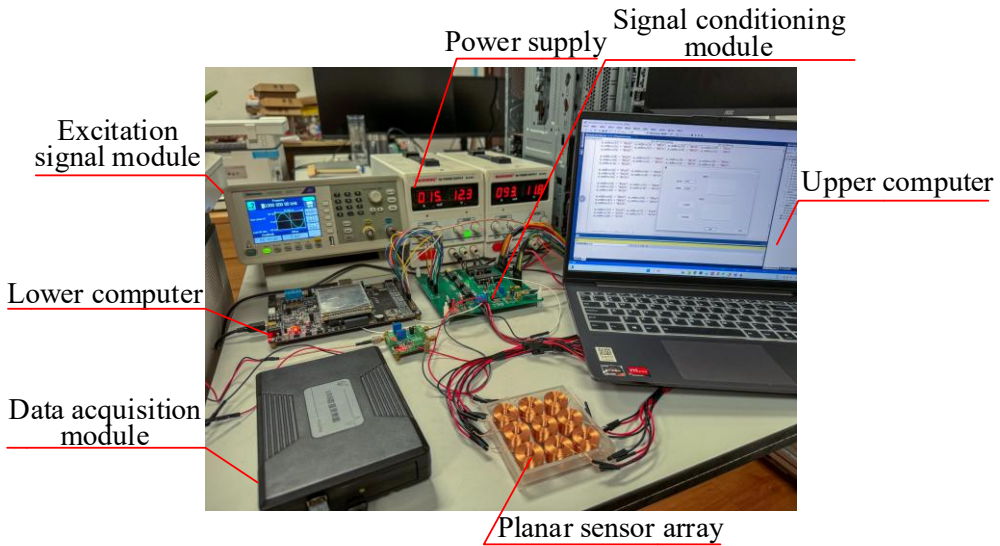


Fig. 12 EMT experimental system of 9 coil planar electromagnetic sensor.

To evaluate the reconstruction performance in real measurement environments, defect detection experiments were conducted using the proposed PI-TransUNet alongside several comparative algorithms. Fig. 13 presents the reconstruction results under various real defect distributions.

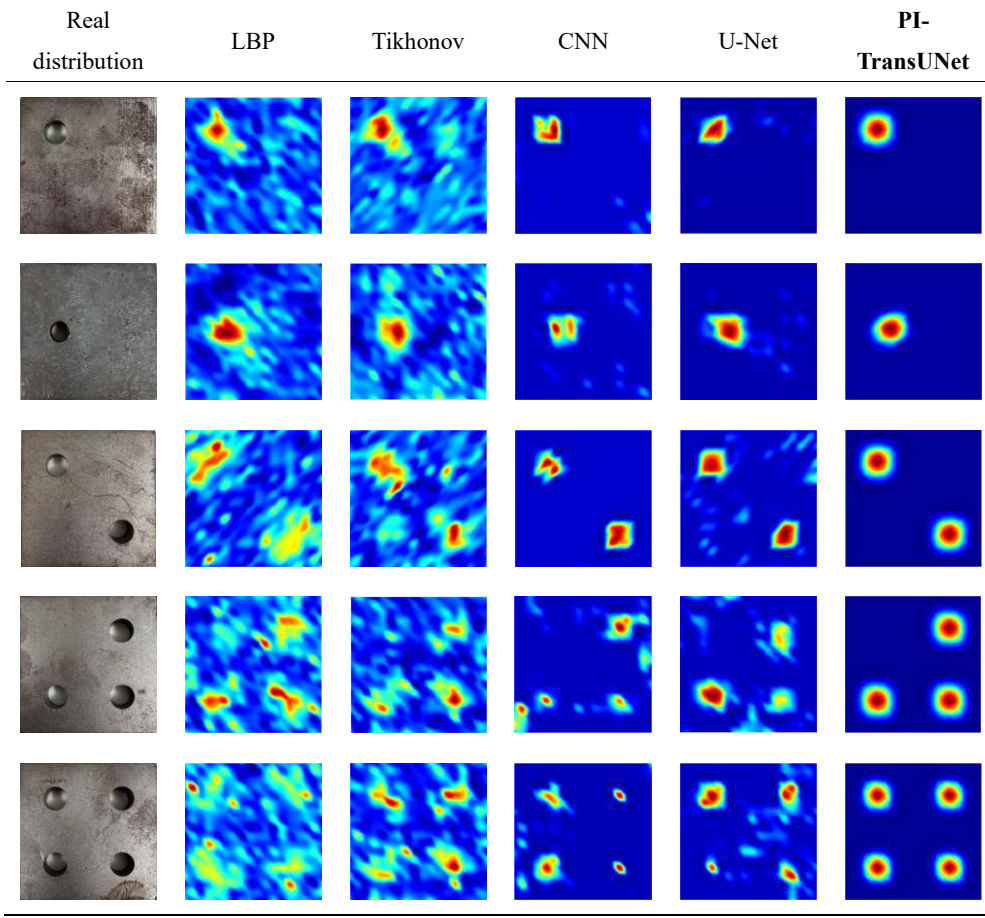


Fig. 13 Experimental reconstruction results obtained with different algorithms.

As shown in Fig. 13, traditional reconstruction methods, including LBP and Tikhonov regularization, suffer from severe background artifacts, blurred boundaries, and poor structural recovery due to measurement noise and the strong ill-posedness of the EMT inverse problem. Although CNN and U-Net markedly suppress background interference, they still exhibit noticeable geometric distortion and degraded performance in complex multi-defect scenarios. By contrast, the proposed PI-TransUNet consistently achieves the best reconstruction performance among all methods. The reconstructed images not only accurately recover the spatial locations of the defects but also preserve their geometric shapes exhibiting strong consistency with the actual defect distributions.

The CC and IE values under five different real defect configurations are summarized in Table 5 to quantitatively evaluate the experimental reconstruction performance. The traditional methods exhibit poor quantitative performance, with CC values generally below 0.16 and IE values exceeding 0.75 in complex defect scenarios. CNN and U-Net achieve improved reconstruction accuracy for simple defects, but degrade significantly as the defect distribution becomes more complicated. For example, in distribution 5, the CC values of CNN and U-Net drop to 0.45 and 0.51, respectively. By comparison, the proposed PI-TransUNet consistently maintains superior reconstruction accuracy and robustness across all experimental conditions, with CC values remaining above 0.82 and IE values below 0.15 for all test real distributions.

Table 5. CC and IE values of experimental reconstruction on the EMT system.

Real distribution	Metrics	LBP	Tikhonov	CNN	U-Net	PI-TransUNet
1	CC	0.3051	0.4151	0.7582	0.8276	0.9215
	IE	0.6124	0.5816	0.2457	0.2186	0.0913
2	CC	0.3514	0.3482	0.6514	0.8345	0.9041
	IE	0.7164	0.7252	0.2577	0.2175	0.1205
3	CC	0.1183	0.1592	0.6455	0.6882	0.8282
	IE	0.8254	0.8683	0.2395	0.2189	0.1473
4	CC	0.1002	0.1520	0.5421	0.7137	0.8695
	IE	0.7692	0.7580	0.6580	0.4475	0.1435
5	CC	0.0912	0.1101	0.4564	0.5174	0.8556
	IE	0.8767	0.8454	0.4513	0.2587	0.1462

Overall, both the visual reconstruction results and quantitative evaluation metrics demonstrate that the proposed PI-TransUNet can achieve high-fidelity and low-error defect imaging in complex real-world EMT detection environments, verifying its strong practical applicability and robustness.

5. Conclusion

This paper has proposed a physics-informed reconstruction framework, termed PI-TransUNet, to address the severe ill-posedness and soft-field effects encountered in the EMT inverse problem for steel plate defect detection. By integrating a differentiable forward physical surrogate model into the end-to-end training process, the proposed method effectively combines data-driven learning with electromagnetic physical constraints. The introduction of spatial positional encoding, together with a global-local hybrid feature extraction strategy, further enables the network to simultaneously capture fine local details and long-range electromagnetic coupling relationships. Extensive simulation and experimental results demonstrate that the proposed PI-TransUNet substantially outperforms both traditional reconstruction methods (LBP and Tikhonov regularization) and conventional deep learning models (CNN and U-Net). Even under severe noise interference (10 dB), the proposed framework maintains strong reconstruction robustness and accurately recovers complex multi-defect structures without noticeable geometric distortion or target adhesion. Moreover, experiments conducted on the custom-built 9-coil planar EMT system verify the strong practical applicability of the proposed method. Across all experimental scenarios, the reconstructed images maintain high structural fidelity, with CC values consistently above 0.82 and IE values below 0.15.

In summary, the proposed PI-TransUNet provides a high-precision, low-artifact, and robust solution for EMT-based non-destructive defect detection. Future work will extend the proposed physics-informed framework to three-dimensional EMT reconstruction and explore multi-parameter joint inversion involving both electrical conductivity and magnetic permeability for complex industrial environments.

CRedit authorship contribution statement

Xianglong Liu: Methodology, Writing – review & editing, Supervision, Project administration, Conceptualization. **Kun Zhang:** Writing – original draft, Software, Methodology, Investigation, Data curation, Formal analysis. **Ying Wang:** Resources, Methodology, Formal analysis. **Huilin Feng:** Formal analysis. **Nan Wang:** Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The work is supported by the National Natural Science Foundation of China (No. 62201511), the Scientific and technological research project in Henan Province (242102220060), and Anhui Province Postdoctoral Research Funding Project (2024C936).

References

- [1] Liu Z, Li W, Xue F, et al. Electromagnetic tomography rail defect inspection [J]. IEEE Transactions on Magnetism, 2015, 51(10): 1-7.
- [2] Wei H Y, Soleimani M. Electromagnetic tomography for medical and industrial applications: Challenges and opportunities [J]. Proceedings of the IEEE, 2013, 101(3): 559-565.
- [3] Xiang J, Dong Y, Yang Y. Multi-frequency electromagnetic tomography for acute stroke detection using frequency-constrained sparse Bayesian learning [J]. IEEE Transactions on medical imaging, 2020, 39(12): 4102-4112.
- [4] Ge L, Liu Z, Liu S, et al. Electromagnetic tomography for multiphase flow in the downhole annulus [J]. IEEE Transactions on Instrumentation and Measurement, 2025.
- [5] Huang P, Huang X, Peng G, et al. Online defect detection on metallic plates using electromagnetic tomography [J]. Insight-Non-Destructive Testing and Condition Monitoring, 2024, 66(2): 109-117.
- [6] Yao J, Takei M. Application of process tomography to multiphase flow measurement in industrial and biomedical fields: A review [J]. IEEE Sensors Journal, 2017, 17(24): 8196-8205.
- [7] Sun B, Yue S, Cui Z, et al. A new linear back projection algorithm to electrical tomography based on measuring data decomposition [J]. Measurement Science and Technology, 2015, 26(12): 125402.
- [8] Liu Z, Yang G, He N, et al. Landweber iterative algorithm based on regularization in electromagnetic tomography for multiphase flow measurement [J]. Flow Measurement and Instrumentation, 2012, 27: 53-58.
- [9] Pastorino M, Caorsi S, Massa A, et al. Reconstruction algorithms for electromagnetic imaging [J]. IEEE Transactions on Instrumentation and measurement, 2004, 53(3): 692-699.
- [10] Xu F, Deshpande M. Iterative nonlinear Tikhonov algorithm with constraints for electromagnetic tomography [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2012, 5(3): 707-716.
- [11] Liu X, Liu Z. A novel algorithm based on L1-Lp norm for inverse problem of electromagnetic tomography [J]. Flow Measurement and Instrumentation, 2019, 65: 318-326.
- [12] Xiang J, Dong Y, Yang Y. Multi-frequency electromagnetic tomography for acute stroke detection using frequency-constrained sparse Bayesian learning [J]. IEEE Transactions on medical imaging, 2020, 39(12): 4102-4112.
- [13] Zhao R, Su Y, Zhang Y, et al. Defect detection of carbon fiber cables by electromagnetic tomography based on proximal gradient method [J]. IEEE Transactions on Instrumentation and Measurement, 2025.
- [14] Li F, Tan C, Dong F, et al. V-net deep imaging method for electrical resistance tomography [J]. IEEE

Sensors Journal, 2020, 20(12): 6460-6469.

- [15] Vu M T, Jardani A. Convolutional neural networks with SegNet architecture applied to three-dimensional tomography of subsurface electrical resistivity: CNN-3D-ERT [J]. Geophysical Journal International, 2021, 225(2): 1319-1331.
- [16] Li X, Liu J, Li Y, et al. ECT image reconstruction algorithm based on Res-SEUnet [J]. Flow Measurement and Instrumentation, 2024, 100: 102688.
- [17] Xiao J, Liu Z, Zhao P, et al. Deep learning image reconstruction simulation for electromagnetic tomography [J]. IEEE Sensors Journal, 2018, 18(8): 3290-3298.
- [18] Liu X, Zhang K, Wang Y, et al. Defect Detection of Steel Plates by Electromagnetic Tomography Imaging Based on Edge-DeepLabv3+ [J]. Journal of Nondestructive Evaluation, 2026, 45(1): 31.
- [19] Liu X, Feng H, Wang Y, et al. Hybrid model of ResNet and transformer for efficient image reconstruction of electromagnetic tomography [J]. Flow Measurement and Instrumentation, 2025, 102: 102843.
- [20] Fu W, Zhang R, Le J, et al. TV-net for 3-D electromagnetic tomography image reconstruction [J]. IEEE Sensors Journal, 2024, 24(21): 34907-34917.
- [21] Liu Y, Zhang R, Fu W, et al. CM2-Net: A Hybrid CNN-Mamba2 Net for 3D Electromagnetic Tomography image reconstruction [J]. IEEE Sensors Journal, 2025.
- [22] Hofmann A, Klein M, Rueter D, et al. A deep residual neural network for image reconstruction in biomedical 3D magnetic induction tomography [J]. Sensors, 2022, 22(20): 7925.
- [23] Jin Y, Li Y, Zhang M, et al. A physics-constrained deep learning-based image reconstruction for electrical capacitance tomography [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 73: 1-12.