(Received 22 December 2025, Accepted 7 January 2026, Scheduled 30 January 2026)

Denoising EMAT Signals and Determining the Thickness of the Sample with a Deep Learning Algorithm

Pan Guo*, Zhuorui Zhang, and Qiuyan Zhong

College of Physics and Electronic Engineering, Chongqing Normal University, Chongqing 401331, China

ABSTRACT: Electromagnetic acoustic transducers (EMATs) have shown broad application prospects in industrial nondestructive testing due to their non-contact and couplant-free operation. However, their low energy conversion efficiency leads to a poor signal-to-noise ratio (SNR), especially under low-power excitation in safety-critical fields such as the petrochemical and nuclear power industries, thereby severely affecting the thickness measurement accuracy. To address this challenge, this paper proposes an Adaptive Dual-Attention Fusion Autoencoder (ADFAE) for EMAT echo signal denoising. The ADFAE adopts a dual-path parallel architecture that integrates a multi-scale convolutional autoencoder with channel attention (MCACA) to capture local temporal features and a spatial attention-guided denoising autoencoder (SAGDA) to model global dependencies. Based on the denoised signals, a CNN-BiLSTM network is further employed to directly estimate material thickness. Experimental results demonstrate that the proposed method achieves effective denoising under low SNR conditions, with an average SNR improvement exceeding 23 dB and a mean Peak SNR above 43 dB. Compared with traditional time-of-flight (TOF)-based methods, the proposed ADFAE-CNN-BiLSTM framework significantly improves thickness measurement accuracy, reducing the average relative error to below 0.25%.

1. INTRODUCTION

Nondestructive testing (NDT) technologies play a critical role in ensuring the structural integrity and operational safety of industrial systems, and they have been widely applied in key sectors such as the aerospace, energy, and petrochemical industries [1, 2]. Among the various NDT techniques, electromagnetic acoustic transducers (EMATs) have emerged as a research focus because of their unique advantages, including non-contact operation, elimination of couplant requirements, suitability for high-temperature environments, and capability for in-line inspection [3, 4]. Compared with conventional piezoelectric ultrasonic testing, EMATs generate and receive ultrasonic waves through electromagnetic coupling, thereby avoiding issues associated with couplants in traditional ultrasonic testing methods [5]. However, this non-contact mechanism also results in significantly lower energy conversion efficiency for EMAT than piezoelectric transducers; consequently, the generated ultrasonic signals exhibit weaker strength and a relatively low signal-to-noise ratio (SNR) [6, 7]. In practical engineering environments within the petroleum, natural gas, and petrochemical industries, equipment is frequently subjected to various sources of electromagnetic interference, vibrational noise, and adverse environmental conditions, which further deteriorate signal quality [8]. Consequently, the echo signals are often completely submerged in background noise, making it difficult to directly determine the material thickness.

To address the problem of low SNR, researchers have proposed various solutions. From a signal processing perspective, traditional methods include signal averaging, digital filtering, and wavelet transform. The variational mode decomposition (VMD) method proposed by Dragomiretskiy and Zosso has been widely applied in the field of signal processing [9]. Si et al. combined VMD with wavelet transform to achieve EMAT signal denoising under large lift-off distance conditions [6]. Guo et al. [10] employed variational mode decomposition for Nuclear Magnetic Resonance echo data denoising, solving the mode mixing and end effect problems of traditional Empirical Mode Decomposition methods. However, these traditional methods require presetting decomposition parameters and lack adaptability, showing limited effectiveness when dealing with complex and variable noise environments. With the rapid development of artificial intelligence technology, deep learning methods have demonstrated tremendous potential in signal processing owing to their powerful feature learning capabilities and noise robustness. Zhou et al. [11] proposed a deep stacked denoising autoencoder for low-power EMAT testing, demonstrating its ability to reconstruct echo signals from voltages ranging from 750 V to 25 V. Using this method, the average SNR was approximately 11 dB at 25 V excitation and thickness measurement relative errors below 0.3% across different specimens. Wu et al. [12] developed a one-dimensional convolutional neural network denoising autoencoder model for pipeline ultrasonic guided wave signal denoising, which improved the SNR by 30.63 dB. Wang et al. [13] performed electrocardiogram signal denoising using a stacked denoising autoencoder, achieving effective noise suppression and robust feature extraction. Lu et al. [14] proposed an Adaptive Denoising Autoencoder for bearing fault vibration signal denoising, achieving an SNR improvement of 24.02 dB and root-mean-square error (RMSE) of 0.13 under varying noise conditions (-6 dB), significantly out-

* Corresponding author: Pan Guo (guopan@cqnu.edu.cn).

performing traditional denoising approaches. However, these methods mostly adopt single-scale feature extraction and lack effective attention mechanisms, making it difficult to adaptively process weak signals in complex noisy environments. Attention mechanisms, as an important breakthrough in the field of deep learning, provide new insights into EMAT signal processing. The Squeeze-and-Excitation networks proposed by Hu et al. [15] significantly improve the image classification accuracy of convolutional neural networks (CNNs) through channel attention mechanisms, which can adaptively adjust the weights of different feature channels and emphasize informative features. The deep residual multi-scale CNN with an attention mechanism proposed by Han et al. [16] significantly improves bearing fault diagnosis accuracy in strong noise environments by adaptively enhancing fault feature representations. These studies indicate that the combination of attention mechanisms and autoencoders has significant potential in EMAT signal processing.

Traditional thickness measurement methods are based on the time-of-flight (TOF) principle of ultrasonic waves in materials [17], which calculates material thickness by measuring the time interval between adjacent echoes and combining it with the ultrasonic wave propagation velocity in the material [18, 19]. This method achieves high accuracy under ideal conditions with good signal quality and homogeneous materials. However, in practical engineering applications, compositional deviations, nonuniform grain sizes, residual stress distributions, and local temperature variations in materials can cause spatial variations in acoustic velocity, thereby affecting the accuracy of thickness determination [20]. Compared to traditional TOF methods, deep learning approaches can directly predict the thickness from raw signals, avoiding the uncertainties of intermediate echo identification steps and providing new insights for thickness determination. Among deep learning technologies, Long Short-Term Memory (LSTM) networks have unique advantages in ultrasonic time-series signal analysis because of their ability to effectively process sequential data [21]. In recent years, hybrid models combining multiple deep learning structures have demonstrated stronger signal processing capabilities. Shang et al. applied a CNN-LSTM hybrid model to ultrasonic guided wave signal processing in metal pipelines, achieving effective identification of six different defect types by extracting 29 time-frequency domain features [22]. Yang et al. [23] proposed a Convolutional Neural Network-Bidirectional Long Short-Term Memory (CNN-BiLSTM) hybrid model for ultrasonic signal thickness evaluation in thin-walled structures, achieving a thickness prediction accuracy of less than 0.05 mm through the extraction of overlapping signal features. The CNN-BiLSTM structure demonstrated superior performance in time-series signal processing by combining the feature extraction capability of CNNs with the sequence modeling capability of BiLSTMs.

Based on the above analysis, this paper proposes an EMAT echo signal denoising and thickness determination method that integrates an Adaptive Dual-Attention Fusion Autoencoder (ADFAE) with a CNN-BiLSTM network. The proposed method aims to address the issue of inaccurate thickness measurement caused by weak signals and strong noise interference in EMAT systems. The main contributions of this work are summarized as follows:

1. An ADFAE is proposed for EMAT echo signal processing, which jointly leverages channel attention and spatial attention mechanisms to achieve effective multi-scale feature extraction and noise suppression.
2. An adaptive feature fusion strategy is introduced within the ADFAE, which dynamically adjusts the contributions of different attention branches based on feature importance, thereby enhancing the model's robustness under varying signal-to-noise ratio conditions.
3. An end-to-end joint framework for EMAT signal denoising and thickness measurement is constructed by integrating the ADFAE with a CNN-BiLSTM regression network, enabling direct prediction from raw echo signals to thickness values and avoiding error accumulation inherent in conventional multi-stage approaches.

The remainder of this paper is organized as follows. Section 2 introduces the EMAT experimental platform and scheme. Section 3 presents the ADFAE-based denoising algorithm and CNN-BiLSTM-based thickness prediction algorithm. Section 4 evaluates the denoising performance and thickness prediction results, and Section 5 provides an overall analysis and discussion of the algorithm performance.

2. DATA ACQUISITION AND SIGNAL PREPROCESSING

2.1. Experimental Setup and Data Acquisition

To validate the proposed method, an EMAT inspection experimental system was established to acquire echo signals. The experimental system consisted of an EMAT device (METP20, ZDJY Company, China), samples of various thicknesses, a data acquisition unit, and a PC, as shown in Fig. 1. The data acquisition computer was connected to the EMAT via Bluetooth and used dedicated software to control thickness measurement parameters and display ultrasonic echo signal waveforms in real time.

The sample used in the experiment was made of 45# steel with dimensions of 60 mm × 60 mm × H mm. The thickness value H, determined with reference to the typical pipe

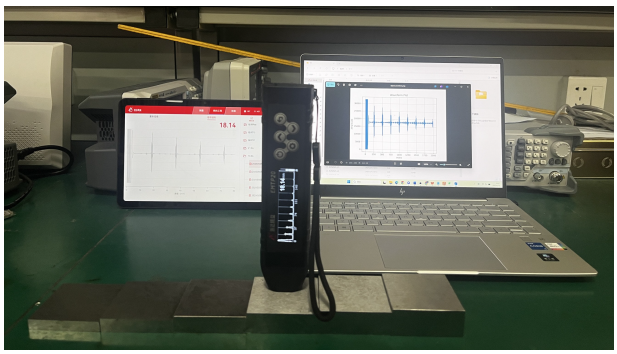

FIGURE 1. EMAT experimental system.

TABLE 1. EMAT raw signal dataset.

Specimen Nominal Thickness (mm)	Actual Thickness (mm)	Training Sample Size	Testing Sample Size
9.0	9.10	80	20
12.0	12.08	80	20
15.0	15.04	80	20
18.0	18.06	80	20
21.0	21.04	80	20
24.0	24.07	80	20

wall thickness ranges specified in API 5L and ASME B36.10 standards and in consideration of engineering practices in the petroleum, natural gas, and chemical industries, was set to six values: 9 mm, 12 mm, 15 mm, 18 mm, 21 mm, and 24 mm. The actual thickness of each specimen was calibrated using a high-precision digital vernier caliper; the calibration data are listed in Table 1. These calibrated thickness values were used as training and testing labels for the deep learning model. To construct a complete training and testing dataset, multi-point measurements were performed on a sample of each thickness to ensure data diversity and representativeness. The EMAT was used to scan clockwise along the upper surface of each specimen, with 100 measurements obtained for each specimen. Upon completion of all scans, a sample database containing 600 A-scan signals was created. The first 150 points of each signal were removed to avoid the influence of the excitation signal. Subsequently, 80% of the data were randomly selected for training, and the remaining 20% were used for independent testing. The original dataset is presented in Table 1.

2.2. Data Augmentation and Noise Simulation

The high cost of EMAT equipment makes experimental data acquisition expensive, requiring substantial material samples and time investment. Moreover, various noise sources in real-world industrial environments significantly affect signal quality, severely compromising thickness measurement accuracy. To approximate real-world application scenarios, this study employs data augmentation techniques to construct training datasets with complex noise interference. A multi-component noise model was established to simulate the signal characteristics in real-world EMAT thickness measurement environments, considering the combined effects of three primary noise sources [24]: (1) electronic noise, modeled as additive white Gaussian noise to simulate random disturbances in circuits; (2) material structural noise, modeled as modulated noise to simulate scattering effects of material microstructures on ultrasonic propagation; (3) environmental background noise, modeled as low-frequency noise to simulate external disturbances such as electromagnetic interference and mechanical vibrations in industrial environments.

This experiment employs 5-fold cross-validation for performance evaluation. Each fold contains 480 training signals, which are partitioned into training and validation sets at a 4:1 ratio. After dataset partitioning, each original EMAT signal undergoes amplitude scaling transformation (scaling factor between 0.8 and 1.2) to simulate variations in the signal under different excitation intensities. Composite noise is then added to generate training samples at six different SNR levels ($-12, -10, -8, -6, -4$, and -2 dB), as shown in Fig. 2. To control noise power, an iterative optimization algorithm is applied to ensure that the actual SNR of each augmented sample deviates no more than ± 0.1 dB from the target value. Ultimately, each fold consists of 2304 training samples and 576 validation samples.

3. METHODS

3.1. Denoising Network Architecture Design

This paper proposes an ADFAE designed to improve denoising and thickness accuracy in EMAT echo signal processing. As shown in Fig. 3, the ADFAE adopts a dual-path parallel processing architecture and incorporates two specially designed denoising modules: a multi-scale convolutional autoencoder with channel attention (MCACA) and a spatial attention-guided denoising autoencoder (SAGDA). The MCACA module captures local temporal features through an optimized encoder-decoder structure, whereas SAGDA models global dependencies to enhance pattern recognition.

The core of this architecture lies in the seamless integration of two complementary attention mechanisms: the channel attention mechanism, which identifies and enhances key frequency components in the signal, and the spatial attention mechanism, which locates and preserves important temporal features. First, the input signal is passed through a multi-scale preprocessing layer for the initial feature decomposition. This layer adopts a dual-branch convolutional structure with kernel sizes of 1×5 and 1×9, respectively, to extract features at different scales. The preprocessed features are then fed into the MCACA module for preliminary denoising. The output of the MCACA is subsequently passed to the SAGDA module for further semantic learning and feature extraction. Finally, the output of SAGDA is fused with the feature representation from the MCACA. To fully exploit the complementary strengths of both modules, the network employs an adaptive weight fusion strategy based on learnable parameters, which is mathemati-

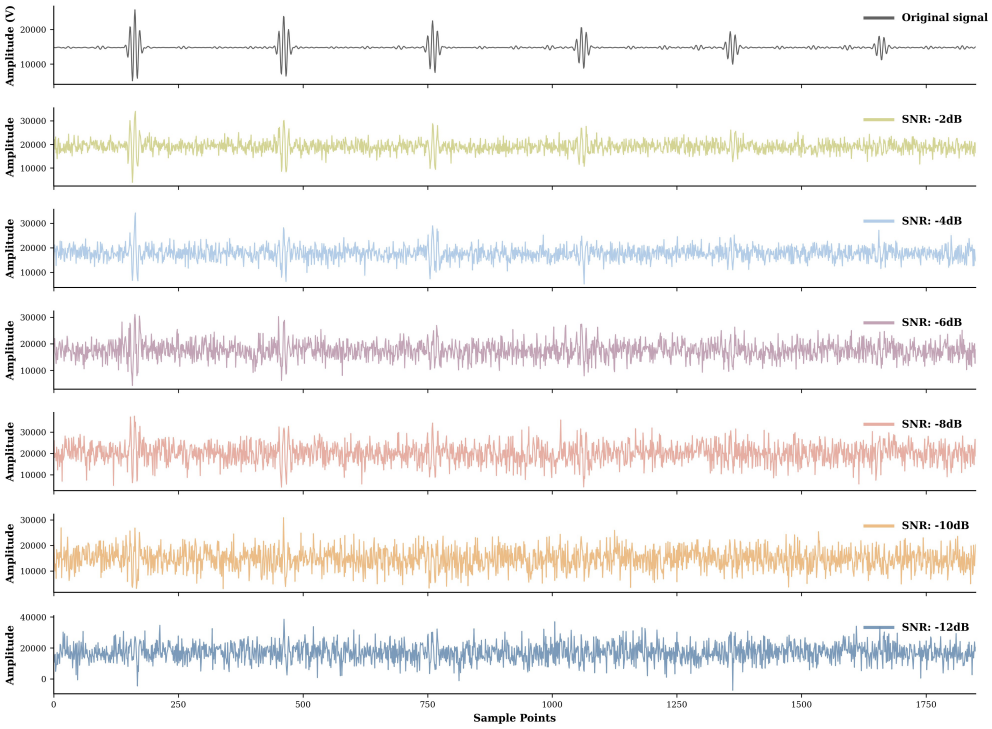

FIGURE 2. EMAT signals with simulated composite noise at different SNR levels.

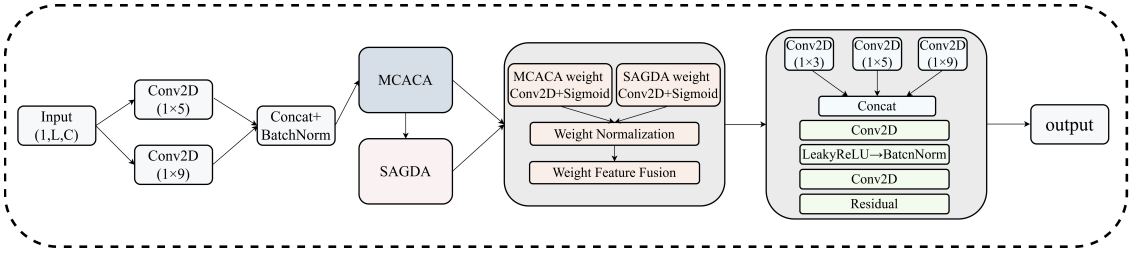

FIGURE 3. Architecture diagram of ADFAE.

cally formulated as:

$$F_{fusion} = \alpha \odot F_{MCACA} + \beta \odot F_{SAGDA} \tag{1}$$

where the weight coefficients are dynamically learned through a normalization operation:

$$\alpha = \frac{W_{MCACA}}{W_{MCACA} + W_{SAGDA} + \epsilon} \tag{2}$$

where F_{MCACA} and F_{SAGDA} represent the feature outputs from the MCACA and SAGDA modules, respectively; F_{fusion} denotes the fused feature representation; α and β are adaptive weight coefficients; $\odot$ denotes element-wise multiplication; W_{MCACA} and W_{SAGDA} are learnable weight parameters generated by the respective weight branches (Conv2D + Sigmoid); and ϵ is a small constant to prevent division by zero. Note that $\beta = 1 - \alpha$ due to the normalization constraint. Finally, high-quality denoised output is generated through multi-scale convolutional branches and residual connections, effectively improving the signal-to-noise ratio and feature fidelity.

3.1.1. MCACA Module Architecture

The MCACA module adopts an improved encoder-decoder structure (as shown in Fig. 4). This module constructs a multi-scale feature extraction mechanism through dual-branch parallel convolution layers, where convolution kernels $k_1 = (1, 5)$ and $k_2 = (1, 9)$ capture local features at different temporal scales. The same stride (1, 1) and padding strategies are employed to maintain the feature map size. In the encoding stage, dilated convolutions with a dilation rate of (1, 2) and a kernel size of (1, 5) are introduced to enhance the modeling capability for long-range dependencies. The decoding process utilizes a symmetric upsampling strategy, and skip connections are used to transfer detailed information, thereby preserving more transient features in the electromagnetic ultrasound signal. Additionally, an enhanced channel attention mechanism is incorporated, which aims to strengthen the channel responses related to key signal features adaptively. Unlike traditional methods that rely solely on global average pooling (GAP), this mechanism integrates global maximum pooling (GMP) to capture

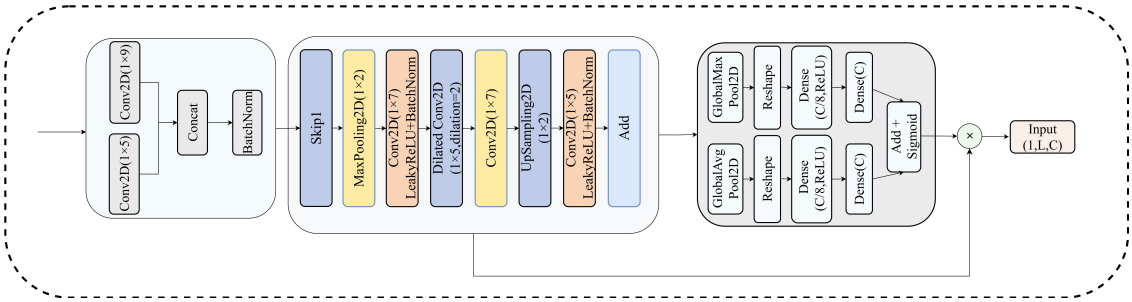

FIGURE 4. Architecture diagram of MCACA.

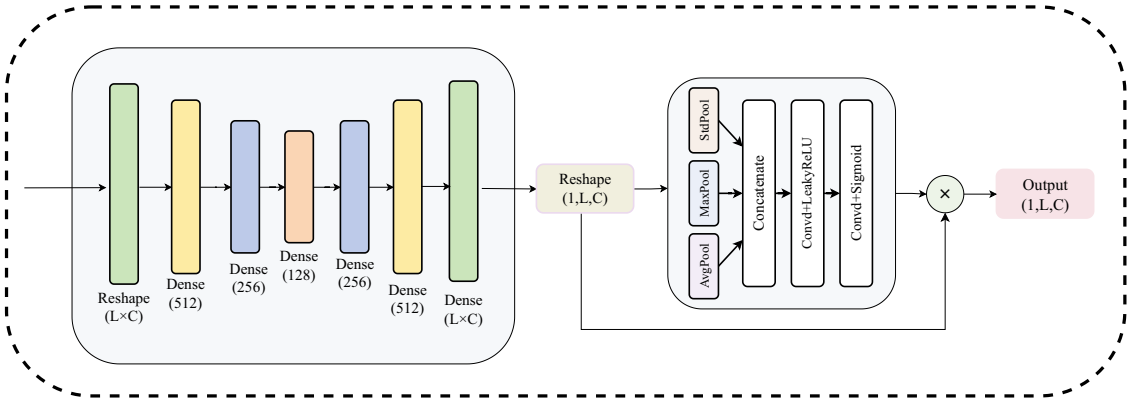

FIGURE 5. Architecture diagram of SAGDA.

richer statistical features. The calculation of channel attention weights follows the following mathematical framework:

$$A_c = \sigma\left(W_2 \cdot \delta\left(W_1 \cdot \mathrm{GAP}(F)\right) + W_2 \cdot \delta\left(W_1 \cdot \mathrm{GMP}(F)\right)\right) \quad (3)$$

where GAP(·) and GMP(·) denote the global average pooling and global maximum pooling operations, respectively; W_1 and W_2 are the shared fully connected layer weight matrices; $\delta(\cdot)$ denotes the ReLU activation function; and $\sigma(\cdot)$ represents the sigmoid activation function. The final feature enhancement is achieved through $F_{enhanced} = A_c \odot F$, where $\odot$ indicates the element-wise multiplication operation.

3.1.2. SAGDA Module Architecture

The SAGDA architecture is depicted in Fig. 5, focusing on modeling global dependencies and complex nonlinear mapping patterns in the EMAT signals. This module adopts a progressive dimensionality reduction strategy, achieving a compact feature representation through a 512→256→128-dimensional encoder, followed by precise reconstruction via a symmetric 256→512→original dimension decoder. Batch normalization and dropout regularization techniques are integrated into each fully connected layer. This module proposes an enhanced spatial attention mechanism. Unlike traditional methods that rely solely on average or max pooling, this mechanism further introduces standard deviation pooling operations to capture higher-order statistical characteristics of signal distributions. Spatial attention uses two-stage convolutions with kernel sizes $k_1 = (1, 7)$ and $k_2 = (1, 3)$, both with the same padding, to generate spatial weight maps. The spatial attention weights are computed as follows:

$$A_s = \sigma\left(\mathrm{Conv}_{(1\times k_2)}\left(\delta\left(\mathrm{Conv}_{(1\times k_1)}(\mathrm{Concat}(F_{avg}, F_{\max}, F_{std}))\right)\right)\right) \quad (4)$$

The definitions of each item are as follows:

$$F_{avg} = \frac{1}{C}\sum_{c=1}^{C} F_c \quad (5)$$

$$F_{\max} = \max_{c\in[1,C]} F_c \quad (6)$$

$$F_{std} = \sqrt{\frac{1}{C}\sum_{c=1}^{C}\left(F_c - F_{avg}\right)^2} \quad (7)$$

where F_{avg}, $F_{\max}$, and F_{std} represent the cross-channel average, maximum, and standard deviation operations, respectively; Concat(·) denotes the feature concatenation operation; and $\mathrm{Conv}_{1\times k}(\cdot)$ represents the two-dimensional convolution transformation with kernel size $(1 \times k)$. The spatial attention mechanism precisely identifies critical temporal positions in the signal through $F_{attended} = A_s \odot F$, significantly enhancing the model's capability for global feature learning and signal reconstruction quality, thereby providing robust support for high-quality denoising of ultrasonic guided wave signals.

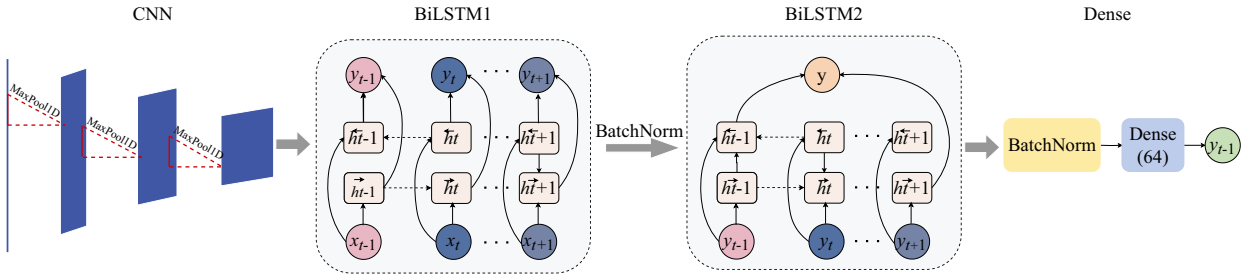

FIGURE 6. Architecture diagram of CNN-BiLSTM.

3.2. Training Configuration

The key hyperparameters, including convolutional kernel sizes, latent dimensions, and loss function coefficients, were determined through preliminary comparative experiments based on validation set performance.

This study employs a hybrid loss function that combines mean squared error (MSE) and gradient terms [25] to ensure signal reconstruction accuracy while preserving the structural characteristics of the signal. The mathematical expression for this hybrid loss function is as follows:

$$\mathcal{L}_{\text{hybrid}} = \alpha \cdot \mathcal{L}_{\text{MSE}} + \beta \cdot \mathcal{L}_{\text{gradient}} \tag{8}$$

where the MSE loss is defined as:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{9}$$

and the gradient preservation loss is defined as:

$$\mathcal{L}_{\text{gradient}} = \frac{1}{N-1} \sum_{i=1}^{N-1} [(y_{i+1} - y_i) - (\hat{y}_{i+1} - \hat{y}_i)]^2 \tag{10}$$

where y_i represents the true signal, $\hat{y}_i$ the reconstructed signal, and N the signal length. The weight coefficients are set as $\alpha = 0.8$ and $\beta = 0.2$ to balance the contributions between reconstruction accuracy and gradient preservation. The weighting coefficients are set as $\alpha = 0.8$ and $\beta = 0.2$ by referring to commonly used configurations in related literature [25] and are verified to yield stable convergence and satisfactory denoising performance in preliminary experiments.

The model is trained using the Adam optimizer with a gradient clipping mechanism, where the clipping threshold is set to 0.5 to prevent gradient explosion. The learning rate scheduling adopts the ReduceLROnPlateau strategy with an initial learning rate of $\eta_0 = 1 \times 10^{-4}$, which monitors the validation loss. When no improvement is observed for 10 consecutive epochs, the learning rate is reduced to half of its current value:

$$\eta_{\text{new}} = 0.5 \times \eta_{\text{current}} \tag{11}$$

The minimum learning rate is limited to $\eta_{\text{min}} = 1 \times 10^{-7}$. This adaptive adjustment mechanism enables the model to converge rapidly during the early training phase while achieving fine parameter tuning through reduced learning rates in later stages. Additionally, an early stopping mechanism is implemented, which terminates training when validation loss shows no improvement for 30 consecutive epochs, effectively preventing overfitting and ensuring the model converges to an optimal solution.

3.3. Thickness Determination Algorithm

In thickness determination tasks, traditional methods generally employ the TOF method [26], which estimates thickness by identifying characteristic peaks in signals and calculating time differences. Ultrasonic waves generate echoes after reflection within the specimen, and the time interval between adjacent echoes is defined as the TOF. Given a constant ultrasonic wave velocity, the specimen thickness can be calculated through the following equation:

$$\text{Thickness} = \text{Velocity} \cdot \frac{\text{TOF}}{2} \tag{12}$$

To further enhance prediction performance, this paper introduces a CNN-BiLSTM hybrid architecture that combines the strengths of CNNs in local feature extraction with the sequence modeling capability of BiLSTM. This design maintains high prediction accuracy under complex noise conditions and provides an efficient and adaptive intelligent solution for electromagnetic acoustic thickness determination.

As shown in Fig. 6, the constructed CNN-BiLSTM model first reshapes the signal into the form of $(n, 1850, 1)$ and inputs it into a three-layer convolutional cascade network. Each convolutional layer uses the same kernel size but employs an increasing number of filters (32, 64, 128), with each convolutional layer followed by a max-pooling layer to achieve signal downsampling and feature compression. The features extracted by the convolutional layers are subsequently fed into the BiLSTM network component, which comprises two BiLSTM layers. The first layer returns all sequences, and the second layer returns only the final state. The final output of the model is calculated as:

$$\begin{aligned}\hat{y} \;=\; & W_o \cdot \text{ReLU}\left(W_h \cdot \text{BiLSTM}_2\left(\text{BiLSTM}_1(\text{CNN}(x))\right)\right. \\ & \left. + b_h\right) + b_o \end{aligned} \tag{13}$$

where $\text{CNN}(x)$ denotes the output of the convolutional layers, and BiLSTM_1 and BiLSTM_2 represent the outputs of the first and second BiLSTM layers, respectively. This hybrid architecture fully leverages the advantages of CNN in local feature extraction and BiLSTM in long-term dependency modeling, enabling the model to process complex one-dimensional signal data more effectively.

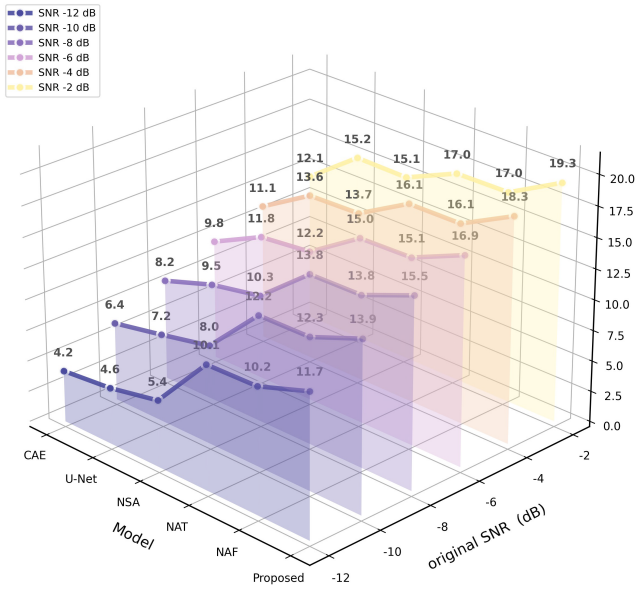

FIGURE 7. Post-denoising SNR of different models under various initial SNR conditions.

4. EXPERIMENTAL RESULTS

4.1. Evaluation Metrics

To comprehensively evaluate the performance of the proposed denoising algorithm, signal-to-noise ratio improvement (SNRI), relative peak position error (RPPE), coefficient of determination (R^2), and peak signal-to-noise ratio (PSNR) [27] are employed as evaluation metrics for assessing the denoising algorithm's performance in denoising tasks. For thickness determination tasks, mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE), and mean relative error (MRE) are adopted to quantify the magnitude and distribution of measurement errors [28]. In Table 2, P_s and P_n represent the power of signal and noise, respectively. SNR_d and SNR_n denote the signal-to-noise ratios of the denoised and noisy signals, respectively. x_i represents the original signal; $\hat{x}_i$ represents the denoised signal; p_i and $\hat{p}_i$ represent the peak positions in the original and denoised signals for the i successfully matched peak pair, respectively; and K is the number of successfully matched peak pairs. $\bar{x}$ denotes the mean value of the original signal; MAX represents the maximum value of the reference signal; y_i is the true thickness value; $\hat{y}_i$ is the predicted thickness value; and n is the number of samples.

TABLE 2. Evaluation metrics.

Denoising Metrics	Thickness Determination Metrics
$\mathrm{SNR}=10\log_{10}\left(\frac{P_s}{P_n}\right)$	$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert$
$\mathrm{SNRI}=\mathrm{SNR}_d-\mathrm{SNR}_n$	$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$
$\mathrm{RPPE}=\frac{1}{K}\sum_{i=1}^{K}\left\lvert\frac{p_i-\hat{p}_i}{p_i}\right\rvert\times 100\%$	$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$
$\mathrm{PSNR}=10\log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right)$	$\mathrm{MRE}=\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i-\hat{y}_i}{y_i}\right\rvert$
$R^2=1-\frac{\sum_{i=1}^{N}(x_i-\hat{x}_i)^2}{\sum_{i=1}^{N}(x_i-\bar{x})^2}$	

4.2. Model Complexity Analysis

To evaluate the computational cost and inference efficiency of the proposed method, the model complexity of each submodule and the overall framework is analyzed in terms of parameter count, model size, and inference time. The ADFAE module contains 91.58 M parameters, with a model size of 349.4 MB, and requires 76.63 ms to process a single EMAT signal. The CNN-BiLSTM module has 1.28 M parameters and a model size of 4.9 MB, with an average inference time of 166.01 ms per signal. Overall, the proposed framework comprises 92.86 M parameters and achieves an average inference time of 242.64 ms per signal, corresponding to a throughput of approximately 4 signals/s. These results indicate that the proposed method is suitable for offline or quasi-real-time EMAT thickness measurement applications.

4.3. Comparison of Denoised Signal Evaluation

To systematically evaluate the effectiveness of the proposed ADFAE model, we select the Convolutional Autoencoder (CAE) [29] and U-Net [30] as baseline models. These architectures are widely used in denoising and feature reconstruction tasks, and are highly relevant to our research problem. The proposed method can be regarded as an improvement and extension of these architectures. Implementing the baselines under identical experimental conditions allows us to demonstrate the performance gains achieved by our method. Moreover, the comparative experiments ensure fair evaluation under the same dataset, data augmentation, and noise conditions, thereby progressively validating the effectiveness of the model. The CAE achieves denoising by learning a compressed representation of the signal through an encoder-decoder architecture, representing a fundamental deep learning approach for signal denoising. On the other hand, U-Net utilizes skip connections to integrate multi-scale features, providing an advanced feature fusion strategy. For the ablation study, we designed variants of the proposed model without SAGDA (NSA), without the attention mechanism (NAT), and without adaptive fusion (NAF) to independently assess the contribution of each core component.

This paper employs a five-fold cross-validation approach, with the overall denoising results evaluation presented in Table 3. The evaluation demonstrates that the proposed model achieves superior performance with an average SNR improvement of 22.93 dB, outperforming other methods by a significant margin, representing a 46.8% improvement over the baseline

TABLE 3. Comparison of overall denoising results across different models.

Model	SNRI (dB)	RPPE (%)	PSNR (dB)	R^2
CAE	15.61±1.80	0.09±0.31	35.68±3.64	0.823±0.130
U-Net	17.31±1.78	0.12±0.35	37.36±4.38	0.859±0.133
NSA	17.81±2.07	0.13±0.37	37.87±4.41	0.878±0.114
NAT	21.02±1.72	0.05±0.11	41.08±2.78	0.952±0.032
NAF	21.08±1.77	0.05±0.10	41.15±2.77	0.953±0.032
Proposed	22.93±2.28	0.04±0.09	43.00±3.52	0.965±0.037

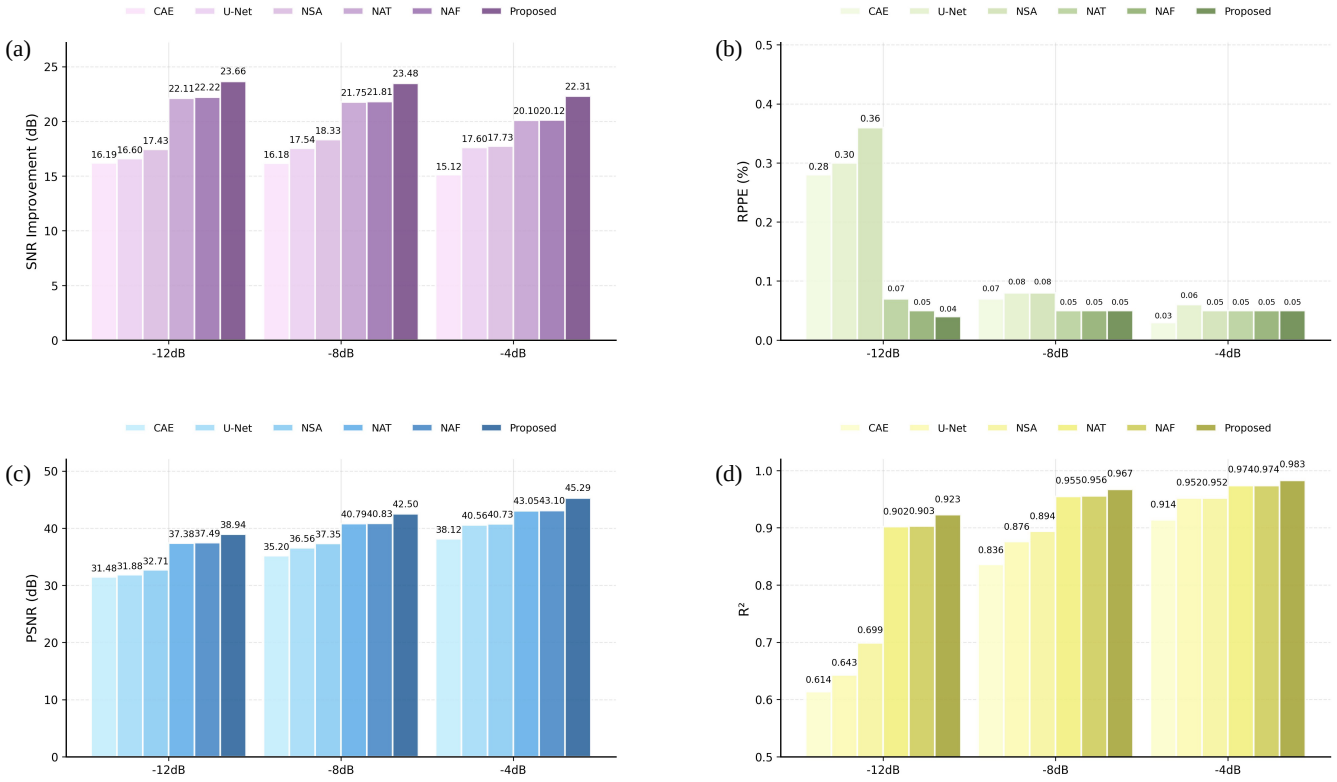

FIGURE 8. Denoising evaluation of different models under three SNR conditions: (a) SNRI, (b) RPPE, (c) PSNR, (d) R^2.

CAE model and a 32.5% improvement over U-Net. In terms of the RPPE metric, the proposed model achieves the lowest value of 0.045%, while CAE and U-Net achieve 0.093% and 0.113%, respectively, demonstrating the superiority of ADFAE in preserving signal peak details. The performance of ablation experiment groups shows a progressive relationship: NSA (17.81 dB) < NAT (21.02 dB) < NAF (21.08 dB), with the contribution of each module clearly visible, where the attention mechanism contributes approximately 1.91 dB improvement, the adaptive fusion strategy contributes approximately 1.85 dB, and the SAGDA module contributes the final 5.12 dB performance enhancement.

Figure 7 intuitively demonstrates the performance gradient variations of each model under different initial SNR conditions. As the initial SNR increases from -12 dB to -2 dB, the post-denoising SNR values of all models exhibit varying degrees of upward trends, with performance differences between models being more pronounced under low SNR conditions, fully demonstrating the advantages of the proposed method in harsh noise environments.

The denoising results of the model under high, medium, and low initial signal-to-noise ratio conditions (-4 dB, -8 dB, -12 dB) are shown in Fig. 8, where the proposed model demonstrates optimal performance across all evaluation metrics. Fig. 8(a) indicates that, in terms of SNR improvement, the proposed model achieves 23.48 dB, 23.89 dB, and 23.66 dB under the three initial SNR conditions, respectively, significantly outperforming other comparative models. The RPPE metric in Fig. 8(b) shows that the proposed model maintains an extremely low level of 0.04–0.05%, demonstrating excellent signal peak position fidelity. The PSNR results (Fig. 8(c)) further validate the model's denoising capability, with the proposed model reaching the highest value of 46.33 dB at -4 dB. The correlation coefficient R^2 achieves the highest values under different initial SNR conditions, indicating high correlation between the denoised signals and the original signals. The ablation study results show that the performance of the three variants, NSA, NAT, and NAF, increases sequentially, confirming the effectiveness of the SAGDA module, attention mechanism, and adaptive fusion strategy.

Figure 9 presents a comparison of denoised signal waveforms from various models under high, medium, and low initial signal-to-noise ratio conditions (-4 dB, -8 dB, -12 dB). Under extremely low SNR conditions (-12 dB), the original signal is almost completely submerged in noise, and while CAE and U-Net can recover the main signal structure, they exhibit obvious waveform distortion and amplitude errors. The NSA model shows limited improvement, whereas the NAT and NAF models demonstrate better waveform preservation capability. The proposed model not only accurately recovers the peak positions and amplitudes of the signal but also maintains the detailed characteristics of the waveform. As the SNR increases to -8 dB and -4 dB, the denoising performance of all models improves; however, the proposed model consistently maintains the best signal reconstruction quality, with its denoised waveforms highly consistent with the original signals, confirming the robustness and adaptability of ADFAE under different noise levels.

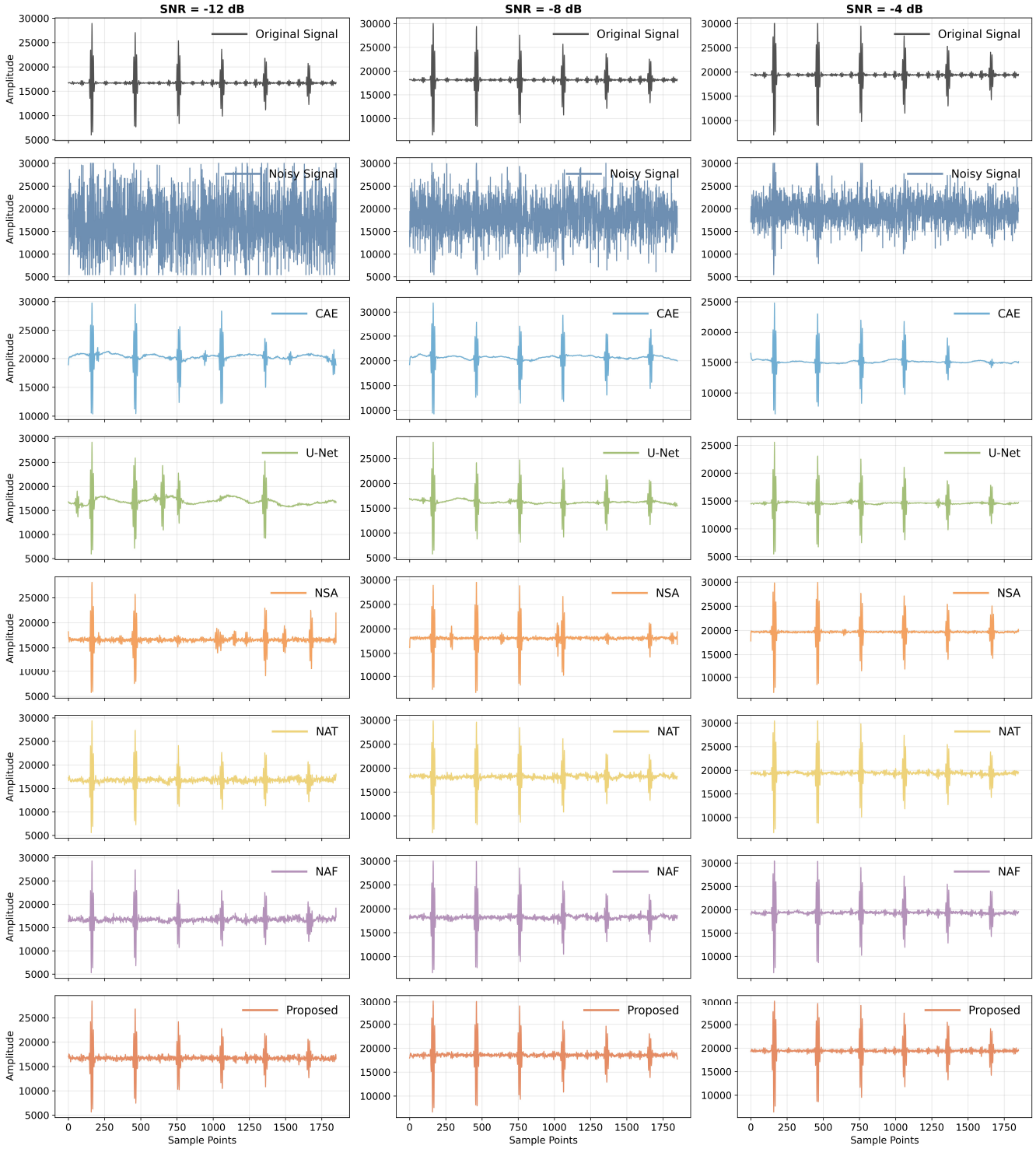

FIGURE 9. Post-denoising SNRs of different models under various initial SNR conditions.

4.4. Evaluation of Thickness Determination Results

After signal denoising, the signals are input into CNN-BiLSTM for thickness determination. This paper compares the overall average regression performance metrics of TOF and CNN-BiLSTM in thickness measurement tasks. The CNN-BiLSTM method employs five-fold cross-validation, and the results are shown in Table 4. The CNN-BiLSTM method significantly outperforms the TOF method across all evaluation metrics. These results indicate that CNN-BiLSTM, by combining the local feature extraction capability of convolutional neural networks with the temporal modeling advantages of bidirectional long short-term memory networks, can more effectively capture complex spatiotemporal features in signals, thereby achieving higher accuracy and more stable performance in thickness prediction tasks.

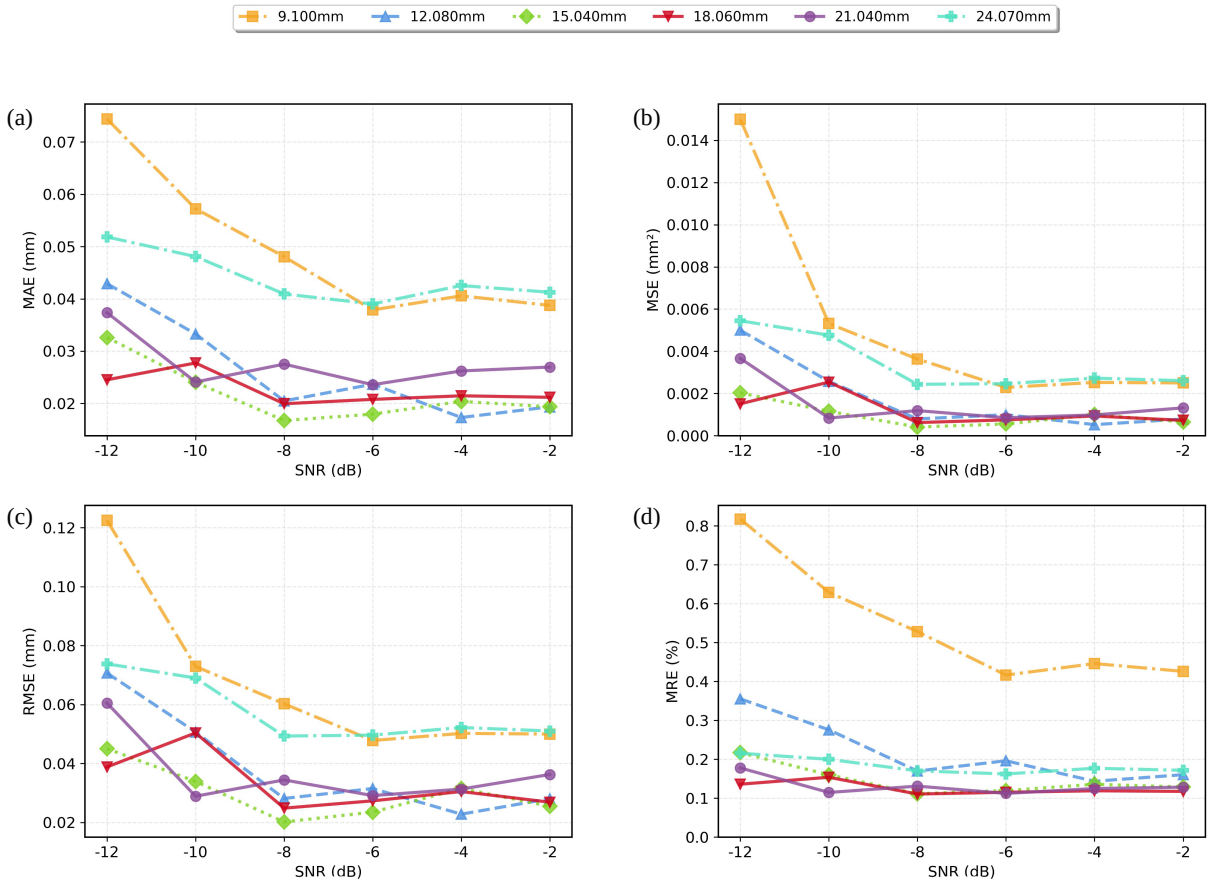

FIGURE 10. Evaluation of thickness determination results under different initial SNR conditions.

TABLE 4. Comparison of thickness determination evaluation between traditional methods and CNN-BiLSTM.

Evaluation Metrics	TOF	CNN-BiLSTM
MAE	0.081	0.044293±0.007192
MSE	0.00765	0.003866±0.002138
RMSE	0.087	0.059992±0.016336
MRE	0.48%	0.2898%±0.0350%

Figure 10 demonstrates the thickness prediction performance of the CNN-BiLSTM hybrid model for six different thickness materials under various signal-to-noise ratio conditions. The results indicate a significant negative correlation between model performance and signal-to-noise ratio. Under low SNR conditions, the 9.100 mm thickness sample exhibits the largest prediction error, with MAE reaching 0.075 mm, MSE of 0.015 mm^2, RMSE of 0.12 mm, and the highest MRE of 0.8%. As the SNR increases, the prediction errors for all thickness samples show a decreasing trend. When SNR ≥ -8 dB, all error metrics tend to stabilize, with MAE stabilizing within the range of 0.02–0.04 mm and MRE decreasing to below 0.2%. Thinner materials exhibit higher prediction errors and greater performance fluctuations in low SNR environments, while medium and thicker materials demonstrate stronger noise immunity, maintaining relatively stable prediction performance under different SNR conditions.

4.5. Comprehensive Application Testing

To evaluate the thickness determination performance of the proposed method, we conducted cascade testing by combining the optimal denoising model and the optimal thickness prediction model from five-fold cross-validation. The experimental test set employed the same data augmentation and noise addition techniques, simulating different initial signal-to-noise ratio conditions of -12 dB, -9 dB, -7 dB, -5 dB, and -2 dB, with each SNR level containing 120 samples. The overall average test results are presented in Table 5, demonstrating excellent performance of the cascade system on the entire test set. The denoising module achieved an SNRI of 23.12 ± 2.20 dB and a PSNR of 43.75 ± 3.28 dB, with a coefficient of determination R^2 reaching 0.9906. The thickness prediction module exhibited even more outstanding performance, with an MRE of 0.25% and a RMSE of 0.042129. Analysis across different SNR conditions reveals that the system maintains stable, high-precision performance under various noise levels.

Based on the heat map analysis results, the system maintains high-precision measurement performance under all test conditions. As shown in Figs. 11(a) and (b), MSE values are distributed within the range of 0.001–0.005 mm^2, and RMSE is controlled between 0.030–0.071 mm, achieving optimal performance under medium thickness conditions (15.040–18.060 mm) and high SNR conditions (-7 dB to -2 dB). Fig. 11(c) shows that MRE exhibits an inverse relationship with thickness, with the thin sample (9.100 mm) showing

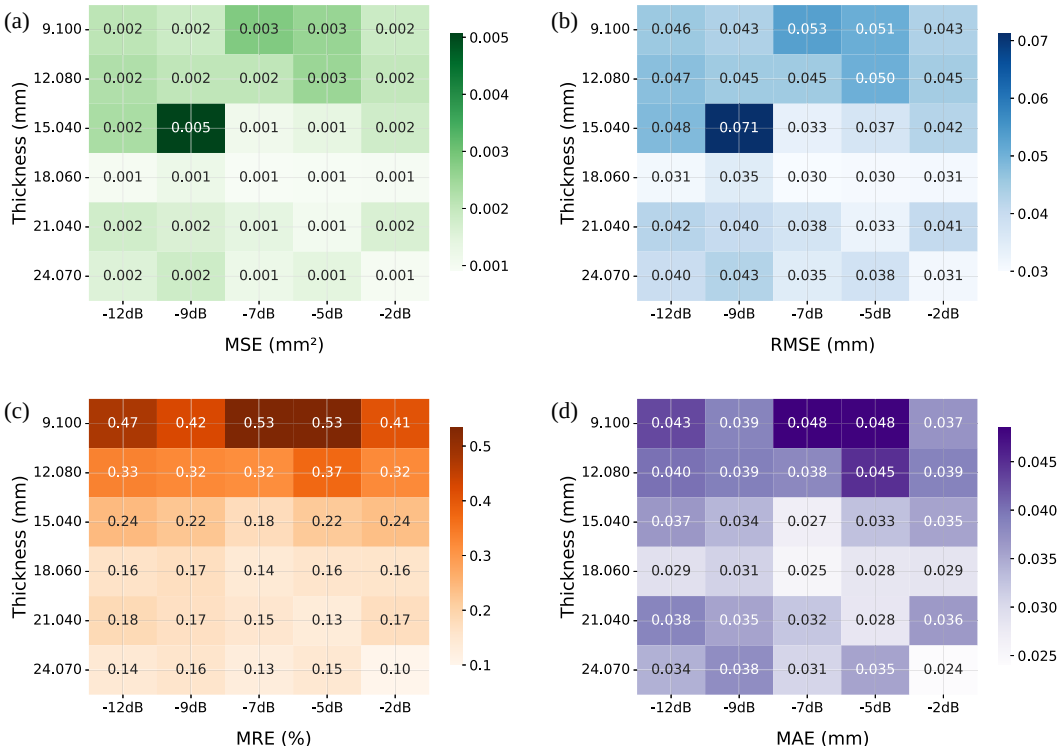

FIGURE 11. Heatmap of thickness determination results under different SNR conditions.

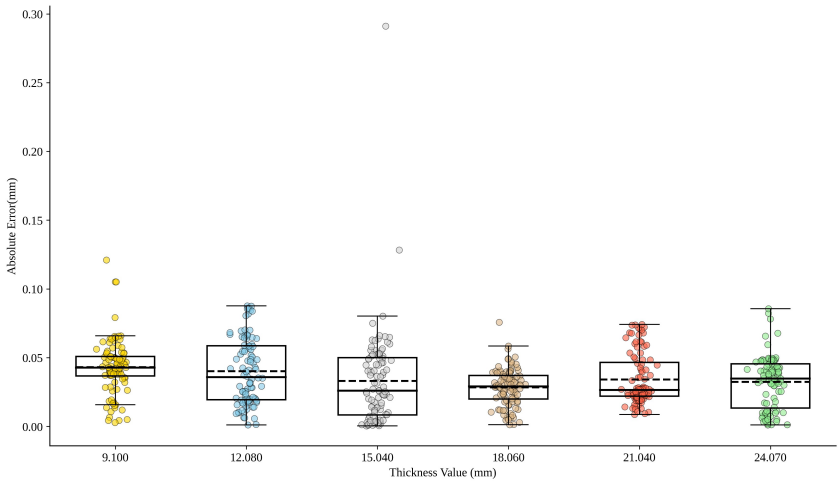

FIGURE 12. Boxplot of overall relative errors on the test set.

TABLE 5. Evaluation of denoising and thickness determination results on the test set.

SNR (dB)	SNRI (dB)	PSNR (dB)	RPPE (%)	R^2	MSE	MAE	MRE (%)	RMSE
Overall	23.12±2.20	43.75±3.28	0.03±0.05	0.9906	0.001775	0.035190	0.25	0.042129
−12	24.46±1.90	39.89±2.30	0.04±0.05	0.9794	0.001830	0.036893	0.26	0.042776
−9	24.22±1.51	42.70±1.95	0.04±0.05	0.9914	0.002272	0.035857	0.24	0.047661
−7	23.11±1.99	43.64±2.21	0.03±0.05	0.9923	0.001585	0.033611	0.24	0.039813
−5	22.49±1.88	44.91±2.26	0.03±0.05	0.9938	0.001649	0.036265	0.26	0.040603
−2	20.88±2.12	46.33±2.35	0.03±0.04	0.9953	0.001539	0.033324	0.23	0.039232

MRE of 0.41%–0.53%, while the thick sample (24.070 mm) demonstrates reduced MRE of 0.10%–0.16%. SNR has a more significant impact on thin samples, with the 9.100 mm sample showing MRE decreasing from 0.53% to 0.41% across

the −12 dB to −2 dB range. Fig. 11(d) indicates that MAE is primarily concentrated within the range of 0.017–0.048 mm, achieving the best absolute accuracy at 18.060 mm thickness with MAE as low as 0.025 mm. The results confirm the stability and effectiveness of the proposed method under different thickness and noise conditions.

Figure 12 presents the absolute errors of all test results, which are primarily distributed within the range of 0.01–0.08 mm. Under the 18.060 mm thickness condition, the system achieves optimal performance with a median error of 0.030 mm, and 75% of the measurement results exhibit errors less than 0.035 mm. Thin samples (9.100 mm) show relatively larger errors with a median of approximately 0.045 mm, while thick samples (21.040–24.070 mm) maintain stable median errors around 0.035 mm. Measurement errors under almost all thickness conditions are controlled within 0.08 mm, with extremely few outliers. This demonstrates that the proposed cascade denoising-thickness determination method can achieve high-precision measurements across different thickness ranges, with optimal performance particularly in the medium thickness range.

5. CONCLUSION

This paper addresses the challenges of weak signals and strong noise leading to low thickness determination accuracy in low-power EMAT used for industrial nondestructive testing. A novel method combining an ADFAE with a CNN-BiLSTM network is proposed for denoising electromagnetic ultrasonic echo signals and calculating material thickness. By introducing the adaptive dual-attention mechanism, the model's denoising capability under low SNR conditions is significantly enhanced, improving the SNR of the denoised signals by more than 23 dB and reducing the RPPE to below 0.05%. Subsequently, the CNN-BiLSTM network predicts the material thickness with a relative measurement error of less than 0.25%, meeting the requirements of safety-critical fields such as petrochemical industries. This approach avoids the conventional reliance on TOF methods for echo peak identification, thereby reducing measurement uncertainty. The method achieves end-to-end thickness measurement without the need for explicit echo recognition.

Despite achieving promising results, several limitations of this study should be acknowledged. First, the validation primarily relies on synthetically noise-augmented data, which may not fully capture the complexity of actual industrial noise environments. Second, the study is conducted solely on 45# steel, and the generalizability to materials with different acoustic properties requires further investigation. Third, as a data-driven approach, the model's reliability for EMAT configurations or thickness values significantly beyond the training range necessitates the supplementation of corresponding training data and subsequent model fine-tuning. Future work will focus on: (1) validation with real industrial data, (2) extension to multiple materials and thicknesses, (3) exploration of the transferability of the approach to other EMAT applications, such as defect detection, and (4) exploration of model compression strategies to enhance inference efficiency in real-time applications.

ACKNOWLEDGEMENT

This work was supported in part by Chongqing Natural Science Foundation (No. cstc2021jcyj-msxmX0470), Technology Funds of Chongqing Municipal Education Commission (KJQN202300524).

REFERENCES

[1] Aramburu, A. B., J. A. da Cruz, A. A. X. da Silva, A. P. Acosta, L. Q. Minillo, and R. de Avila Delucis, "Non-destructive testing techniques for pressure vessels manufactured with polymer composite materials: A systematic review," *Measurement*, Vol. 246, 116729, 2025.

[2] Boughedda, H., T. Hacib, Y. L. Bihan, and H. Acikgoz, "Cracks characterization of non-ferromagnetic material using emat probe and PLSR technique," *Progress In Electromagnetics Research C*, Vol. 103, 199–209, 2020.

[3] Huang, X., Y. Xie, F. Liu, and L. Xu, "Electromagnetic acoustic damage imaging using compressive sensing and total focusing method," *IEEE Transactions on Instrumentation and Measurement*, Vol. 74, 1–9, 2025.

[4] Wang, C., J. Ma, X. Bai, and J. Chen, "Defect detection and imaging using electromagnetic acoustic transducer with butterfly coil," *Review of Scientific Instruments*, Vol. 95, No. 6, 064704, 2024.

[5] Shankar, S. and K. Balasubramaniam, "Characterising the beam formation of SH waves using double-row staggered Halbach EMAT configurations," *NDT & E International*, Vol. 121, 102465, 2021.

[6] Si, D., B. Gao, W. Guo, Y. Yan, G. Y. Tian, and Y. Yin, "Variational mode decomposition linked wavelet method for EMAT denoise with large lift-off effect," *NDT & E International*, Vol. 107, 102149, 2019.

[7] Cheng, L., M. Kogia, A. Mohimi, V. Kappatos, C. Selcuk, and T.-H. Gan, "Crack characterisation using invariable feature extraction in stainless steel specimen used for absorber tubes of CSP applications via EMAT," *Renewable Energy*, Vol. 101, 771–781, 2017.

[8] Mudapaka, V., T. Gantala, and K. Balasubramaniam, "Development of bulk wave EMAT sensors with enhanced Lorentz force through magnetic field concentration in eddy current regions," *Ultrasonics*, Vol. 154, 107687, 2025.

[9] Dragomiretskiy, K. and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, Vol. 62, No. 3, 531–544, 2014.

[10] Guo, J., R. Xie, Y. Wang, L. Xiao, J. Fu, G. Jin, and S. Luo, "Variational mode decomposition for NMR echo data denoising," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 61, 1–14, 2023.

[11] Zhou, J., D. Yu, X. Li, Y. Zheng, and Y. Liu, "Reconstructing echoes completely submerged in background noise by a stacked denoising autoencoder method for low-power EMAT testing," *Measurement Science and Technology*, Vol. 34, No. 12, 125910, Sep. 2023.

[12] Wu, J., Y. Yang, Z. Lin, Y. Lin, Y. Wang, W. Zhang, and H. Ma, "Weak ultrasonic guided wave signal recognition based on one-dimensional convolutional neural network denoising autoencoder and its application to small defect detection in pipelines," *Measurement*, Vol. 242, 116234, 2025.

[13] Wang, E. K., X. Zhang, and L. Pan, "Automatic classification of CAD ECG signals with SDAE and bidirectional long short-term network," *IEEE Access*, Vol. 7, 182 873–182 880, 2019.

[14] Lu, H., K. Zhou, and L. He, "Bearing fault vibration signal denoising based on adaptive denoising autoencoder," *Electronics*, Vol. 13, No. 12, 2403, 2024.

[15] Hu, J., L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7132–7141, Salt Lake City, UT, USA, 2018.

[16] Han, S., S. Sun, Z. Zhao, Z. Luan, and P. Niu, "Deep residual multiscale convolutional neural network with attention mechanism for bearing fault diagnosis under strong noise environment," *IEEE Sensors Journal*, Vol. 24, No. 6, 9073–9081, 2024.

[17] Zhang, Y., S. L. Huang, S. Wang, and W. Zhao, "Time-frequency energy density precipitation method for time-of-flight extraction of narrowband lamb wave detection signals," *Review of Scientific Instruments*, Vol. 87, No. 5, 054702, 2016.

[18] Liu, W., D. Li, and W. Xie, "A novel time-of-flight difference determination method for ultrasonic thickness measurement with ultrasonic echo onset point detection," *Applied Acoustics*, Vol. 233, 110605, 2025.

[19] Dive, V. P., S. Solepatil, A. Mali, S. Dambhare, and A. Kolekar, "Investigation of the effect of grain size on longitudinal critically refracted ultrasonic wave time-of-flight and velocity of propagation in material," *Advances in Science and Technology Research Journal*, Vol. 19, No. 6, 391–402, 2025.

[20] Wan, T., T. Naoe, T. Wakui, M. Futakawa, H. Obayashi, and T. Sasa, "Effects of grain size on ultrasonic attenuation in type 316L stainless steel," *Materials*, Vol. 10, No. 7, 753, 2017.

[21] Zhou, Y. and X. Jiao, "Intelligent analysis system for signal processing tasks based on LSTM recurrent neural network algorithm," *Neural Computing and Applications*, Vol. 34, No. 15, 12 257–12 269, 2022.

[22] Shang, L., Z. Zhang, F. Tang, Q. Cao, H. Pan, and Z. Lin, "CNN-LSTM hybrid model to promote signal processing of ultrasonic guided lamb waves for damage detection in metallic pipelines," *Sensors*, Vol. 23, No. 16, 7059, 2023.

[23] Yang, L., J. Zhao, L. Chang, J. Zhao, and J. Cheng, "Evaluating the thickness of thin-walled structures from ultrasound signals based on CNN-BILSTM model," *Nondestructive Testing and Evaluation*, 1–19, 2025.

[24] Kubinyi, M., O. Kreibich, J. Neuzil, and R. Smid, "EMAT noise suppression using information fusion in stationary wavelet packets," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 58, No. 5, 1027–1036, 2011.

[25] Lu, Z. and Y. Chen, "Single image super-resolution based on a modified U-Net with mixed gradient loss," *Signal, Image and Video Processing*, Vol. 16, No. 5, 1143–1151, 2022.

[26] Maev, R. G., H. Shao, and E. Y. Maeva, "Thickness measurement of a curved multilayered polymer system by using an ultrasonic pulse-echo method," *Materials Characterization*, Vol. 41, No. 2-3, 97–105, 1998.

[27] Fan, C., Y. Zhang, H. Ma, and Z. Ma, "A novel deep denoising model integrating transformer and time-frequency loss for gearbox fault diagnosis," *Advanced Engineering Informatics*, Vol. 66, 103400, 2025.

[28] Wang, P., H. Tao, J. Qi, and P. Li, "Machining quality prediction of complex thin-walled parts using multi-task dual domain adaptive deep transfer learning," *Advanced Engineering Informatics*, Vol. 62, 102640, 2024.

[29] Lomoio, U., P. Veltri, P. H. Guzzi, and P. Liò, "Design and use of a denoising convolutional autoencoder for reconstructing electrocardiogram signals at super resolution," *Artificial Intelligence in Medicine*, Vol. 160, 103058, 2025.

[30] Cai, J., L. Wang, J. Zheng, Z. Duan, L. Li, and N. Chen, "Denoising method for seismic co-band noise based on a U-Net network combined with a residual dense block," *Applied Sciences*, Vol. 13, No. 3, 1324, 2023.