# Intelligent Identification of Pile Defects Based on Improved LSTM Model and Wavelet Packet Local Peaking Method

Xiaolin Li[1], Xinyi Chen[2]*
Anhui Construction Engineering Quality Supervision and Inspection Station Co., Ltd, Hefei, China[1]
Anhui and Huaihe River Institute of Hydraulic Research, Hefei 230000, China[1]
Anhui Vocational and Technical College, School of Civil Engineering, Hefei 230011, China[2]

***Abstract*—With the continuous expansion of building scale, the structural safety of foundation piles, as key load-bearing components, has received increasing attention. To improve the defect recognition ability under complex working conditions, this study first uses the whale optimization algorithm to perform hyperparameter optimization on the long short-term memory network model, achieving efficient classification of the defect and non-defect samples. Subsequently, the signals identified as having defects are subjected to wavelet packet decomposition to extract multi-scale energy features, and combined with the local peak finding method to accurately locate key reflection peaks, achieving further identification of defect types. The results showed that the classification accuracy, recognition precision, recall rate, and F1 value of the new method were the highest at 96.7%, 95.16%, 93.87%, and 94.51%, respectively, and the average recognition time was the shortest at 0.97 seconds. Especially for the defect identification errors of drilled cast-in-place piles and prefabricated piles, the lowest were 0.19 and 0.23, and the lowest complexity could reach 65.28%, demonstrating high precision and stability in defect identification. This model has strong robustness and accuracy in various types of defect scenarios, and has good generalization ability and engineering application potential, which can provide certain technical references for the construction monitoring of road and bridge engineering in the future.**

***Keywords*—Foundation pile; defect identification; LSTM; WOA; WPT; LPS**

## I. Introduction

As the core of the building foundation, the quality of foundation piles directly determines the stability and safety of the building. Especially in areas with complex geological conditions, there is a high risk of potential defects in foundation piles. If these defects are not detected and treated promptly, they may pose a serious threat to the overall safety of buildings [1-2]. Therefore, early identification and accurate evaluation of pile defects have become key issues that urgently need to be addressed in current construction quality management. With the development of artificial intelligence technology, machine learning-based intelligent recognition methods have gradually become an important research direction in pile defect detection. Wu J et al. developed a multi-point traveling wave decomposition method for detecting and characterizing damage in cast-in-place reinforced concrete piles. This method was more effective in extracting damage features and had higher recognition and detection accuracy compared to other advanced methods [3]. Zhang W et al. developed a novel detection way grounded on the image segmentation network U-Net. Compared with traditional algorithms, the developed algorithm exhibited better performance in terms of accuracy and F1 value [4]. Jiang S et al. proposed an underwater pile defect detection model by combining an image fusion enhancement algorithm and a deep learning algorithm. This model had good robustness to noise and performed well in surface defect detection [5]. Liu H et al. put forth a new non-destructive testing method to solve the detection problems of concrete disintegration or steel corrosion [6]. This method could achieve high-frequency identification of defects in sensitive areas of pile foundations but required high detection conditions.

Deep learning algorithms, especially Long Short-Term Memory (LSTM) networks, have been widely used in various intelligence recognition tasks due to their advantages in time series data analysis and pattern recognition [7]. Wu C S et al. believed that the efficiency of utilizing conventional methods to identify multiple kinds of defects in pile foundations was very low, and proposed a pile-based defect type identification method built on dual channel Convolutional Neural Networks (CNN) and LSTM. It effectively integrated 1D and 2D features, extracted more potential features, and improved classification precision [8]. Wang H et al. proposed a low-strain pile foundation detection data method based on recursive neural networks and improved LSTM. In comparison, this method had the highest accuracy but required more training parameters [9]. Wu J et al. proposed a new multi-sensor pile damage detection method that can effectively identify damage in a multi-task learning framework [10]. Hu T et al. developed a new detection method based on improved LSTM to address the shortcomings of existing methods for predicting the settlement of surrounding buildings caused by deep excavation construction [11]. The settlement predicted by this method under three working conditions was in good agreement with the monitored settlement.

In summary, although existing studies have made some progress in the recognition of foundation pile defects, most of them focus on a single model and suffer from the problems of decoupling of classification and localization, sensitivity to high-frequency noise, and insufficient expression of defect features. LSTM is chosen as the fundamental model for this study. The reason is that LSTM has excellent time-dependent modeling capabilities, suitable for processing time series features of complex pile detection signals, and can effectively alleviate the gradient vanishing problem of traditional recurrent neural networks. However, the LSTM model is sensitive to hyperparameter settings and is deficient in multi-scale defect feature capture and localization accuracy when used alone. To compensate for these shortcomings, the study introduces the Whale Optimization Algorithm (WOA) to globally optimize the hyperparameters of the LSTM, which improves the robustness and generalization ability of the model. At the same time, it combines the Wavelet Packet Transform-Local Peak Search (WPT-LPT) to optimize the LSTM hyperparameters and enhance the multi-scale energy decomposition of defective signals and the localization of key peaks, forming a synergistic identification system. The method aims to overcome the limitations of existing models and realize high-precision classification and localization of foundation pile defects. The method innovatively uses WOA for LSTM hyperparameter global optimization, which effectively improves the robustness and generalization ability of the classification model. WPT is used to realize multi-scale energy decomposition of the signal, and combined with the LPS recognition strategy, it enhances the ability to respond to defective mutation features. Different from the traditional single-stage identification method, the model establishes a joint identification process of "classification-decomposition-localization", which can simultaneously output structured information such as the existence and type of defects. The research method provides high-precision, low-latency, and robust solution for monitoring the health of pile foundations under complex working conditions.

The study is divided into four sections: Section I introduces the background of foundation defect identification and the improved LSTM classification method based on WOA optimization. Section II describes the multi-scale signal feature extraction and defect localization by combining the WPT and the LPS methods. Section III carries out the model performance test and the ablation analysis to validate its accuracy and robustness. Section IV concludes the results of the study, points out the limitations, and looks forward to future applications.

## II. Methods and Materials

### A. Classification of Pile Defects Based on Improved LSTM

In the actual construction and service process, various factors such as geological conditions, construction techniques, and material quality may affect the occurrence of different types of defects in pile structures [12-13]. These defects may not only weaken the bearing capacity of foundation piles but also cause settlement, tilting, and even overall structural damage during long-term service, and in severe cases, can lead to catastrophic failure of bridge structures [14-15]. Fig. 1 shows the common forms of pile defects.

In Fig. 1, structural defects mainly include fractures, shrinkage, and displacement, which are often caused by abnormal stress or changes in geological conditions. Material defects such as insufficient strength, honeycomb, and rough surface are usually closely related to the quality of the concrete itself. Construction process defects reflect human factors in the pile foundation construction process, such as incorrect positioning of steel bars and incomplete hole cleaning. Various types of defects may appear individually or in combination in engineering. Therefore, higher requirements have been put forward for detection and recognition. Therefore, introducing deep learning methods with temporal modeling capabilities has become an effective path to improve defect recognition performance. LSTM, as an improved structure of Recurrent Neural Networks (RNNs), has a strong time-series learning ability and can avoid the gradient vanishing problem that traditional RNNs encounter when the sequence is long while maintaining long-term context-dependent information [16-18]. Fig. 2 shows the structure of a stacked LSTM.

In Fig. 2, the stacked LSTM is still composed of the basic LSTM, with intermediate connecting layers. Its basic unit consists of three gate control structures, namely Forget Gate (FG), Input Gate (IG), and Output Gate (OG). At each time step $t$, a single LSTM unit dynamically regulates the information flow through three gating mechanisms. The IG determines which new information is introduced into the memory unit at the current time based on the current input $x_t$ and the hidden state $h_{t-1}$ of the previous time. The calculation formula is shown in Eq. (1):

$$\begin{cases} i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \end{cases} \tag{1}$$

In Eq. (1), $i_t$ is the output vector of the IG. $\sigma$ is the Sigmoid activation function. $W_i$ and $W_C$ are weight matrices for IGs and candidate states. $h_{t-1}$ is the hidden state of the previous time step. $x_t$ is the input vector for the current time step. $b_i$ and $b_C$ are bias vectors for IGs and memory states. $\tilde{C}_t$ is a candidate memory state. The expressions for the FG and OG are shown in Eq. (2):

$$\begin{cases} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t = o_t * tanh(C_t) \end{cases} \tag{2}$$

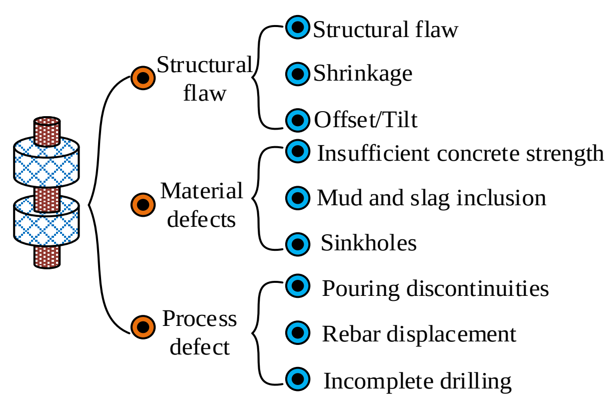



Fig. 1. Common defect form of foundation pile

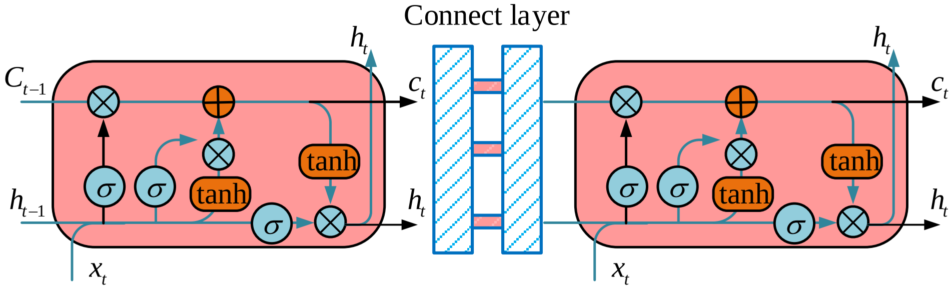



Fig. 2. Stacked LSTM structure

In Eq. (2), $f_t$ and $o_t$ are the output values of the FG and OG. $W_f$ and $W_o$ are the weight matrices of the FG and OG. $b_f$ and $b_o$ are bias vectors for the FG and OG. $C_t$ is the current state of the memory unit, as shown in Eq. (3):

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{3}$$

Although LSTM has shown good classification ability in temporal modeling, its performance is highly dependent on the setting of model hyperparameters. In the task of identifying pile foundation defect images or signals, the problems of complex feature distribution and imbalanced samples are commonly present. Traditional manual parameter adjustment methods are not only inefficient, but also prone to falling into local optima. Therefore, this study introduces WOA for parameter optimization of LSTM model. Firstly, in each round of optimization, WOA guides individuals in the population towards position $\vec{X}(t)$ based on the current optimal parameter combination position $\vec{X}^*(t)$, and the position update is shown in Eq. (4):

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A}\cdot\left|\vec{C}\cdot\vec{X}^*(t) - \vec{X}(t)\right| \tag{4}$$

In Eq. (4), $\vec{A} = 2\vec{a}\cdot\vec{r}_1 - \vec{a}$, $\vec{C} = 2\cdot\vec{r}_2$, where $\vec{r}_1$ and $\vec{r}_2$ both represent random vectors in the interval [0,1]. $\vec{a}$ is the control factor. To enhance local search capability, WOA introduces a spiral approximation mechanism to simulate the nonlinear convergence path of whales around prey, as calculated in Eq. (5):

$$\vec{X}(t+1) = e^{bl}\cdot\cos(2\pi l)\cdot\left|\vec{X}^*(t) - \vec{X}(t)\right| + \vec{X}^*(t) \tag{5}$$

In Eq. (5), $b$ is the helical contraction factor. $l$ is a random number within the range of [-1,1]. Finally, after each iteration, WOA evaluates the individual fitness and dynamically updates the global optimal parameter combination based on the current optimal fitness value, as shown in Eq. (6):

$$\iota = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{ik}\log(\hat{y}_{ik}) \tag{6}$$

In Eq. (6), $N$ is the gross of samples. $K$ is the amount of defect categories. $y_{ik}$ is the true label of sample $i$ belonging to class $k$. $\hat{y}_{ik}$ is the prediction probability of the LSTM model for class $k$. The pile defect classification process of WOA-LSTM is shown in Fig. 3.

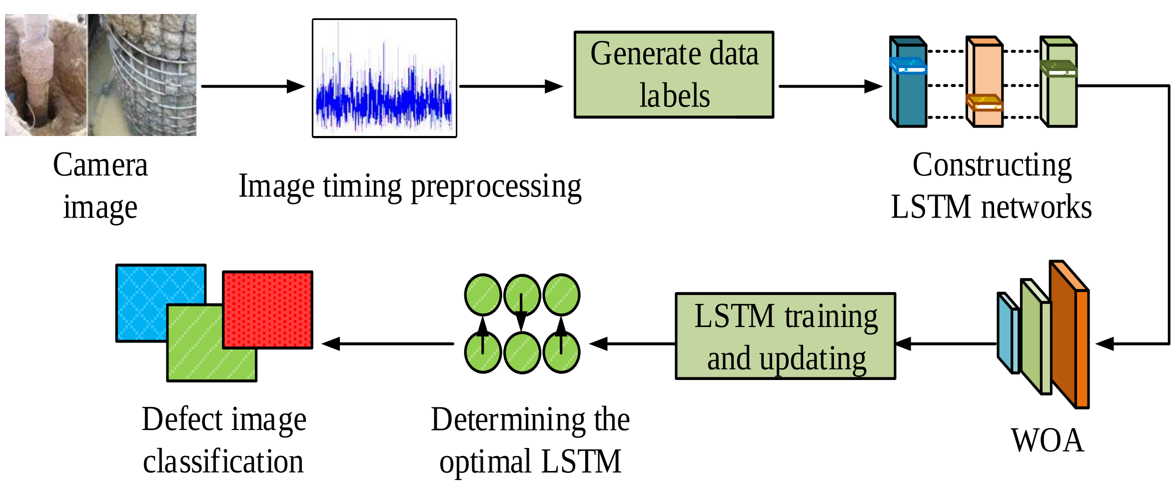



Fig. 3. Defect classification process of foundation pile based on WOA-LSTM

In Fig. 3, firstly, the system performs preprocessing operations such as cleaning and normalization on the collected pile foundation detection signals or defect images to generate defect category labels. Secondly, an initial LSTM model is constructed, and multiple candidate hyperparameter combinations are initialized through WOA to form a search population. Next, the LSTM model will be trained for each set of parameters, and performance evaluation will be conducted using the cross entropy loss function as the fitness metric. Subsequently, WOA performs position updates and spiral local search operations, guiding the population to continuously approach the optimal solution. After the iteration, a set of globally optimal LSTM hyperparameter combinations is obtained, and the model is retrained based on this to achieve higher accuracy in defect classification. Finally, the optimized model is used to identify defect types in newly collected signal or image data.

### B. *Construction of PDI Model Integrating WPT-LPS Method*

After completing the construction of the WOA-LSTM-based pile defect classification process, further research finds that the data obtained in actual pile foundation testing usually has significant non-stationarity and high noise interference. Especially reflected wave signals, sound wave transmission images, etc., these information often contain important features of defects. Therefore, this study introduces WPT as a front-end preprocessing technique to perform multi-level decomposition on the original signal. Compared with ordinary wavelet transform, wavelet packets can simultaneously decompose the high-frequency and low-frequency parts of the signal into full-frequency bands and have stronger time-frequency localization ability [19-20]. The WPT signal decomposition diagram is shown in Fig. 4.

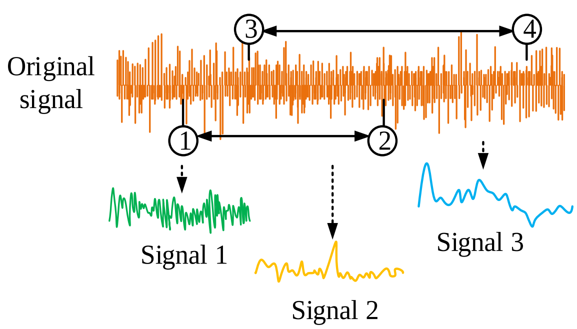



Fig. 4. Diagram of WPT signal decomposition.

In Fig. 4, WPT performs recursive full-frequency convolution decomposition on the original pile foundation detection signal, refining the original non-stationary signal into multiple independent sub-signals in various frequency bands. The decomposition process can be seen as a hierarchical filtering downsampling operation. Each level decomposes the signal from the previous level into two new sub-signals, namely a low-frequency component and a high-frequency component. The formula for signal convolution decomposition is shown in Eq. (7):

$$W_{n+1,2g}(t)=\sum_{z}x_n(z)\cdot h(2t-z) \tag{7}$$

In Eq. (7), $W_{n+1,2g}(t)$ is the signal of the $2g$ -th node in the $n$ -th layer. $x_n(z)$ is the input signal of the $z$ -th node in the $n$ -th layer. $h$ is the low-pass filter coefficient of WPT. The calculation for obtaining high-frequency sub bands is shown in Eq. (8):

$$W_{n+1,2g+1}(t)=\sum_{z}x_n(z)\cdot v(2t-z) \tag{8}$$

In Eq. (8), $W_{n+1,2g+1}(t)$ is the signal of the $2g+1$-th node in the $n$ -th layer. $v$ is the coefficient of the high pass filter. Eq. (7) and (8) show that low-frequency sub-signals reflect a steady trend, while high-frequency sub-signals typically contain sensitive responses to defect mutations. The energy calculation of a single node is shown in Eq. (9):

$$E_{n,g}=\sum_{t}\left|W_{n,g}(t)\right|^2 \tag{9}$$

In Eq. (9), $E_{n,g}$ is the energy value of node $(n,g)$. To more accurately extract key features related to defects from multi-layered sub-bands, this study further introduces LPS to identify mutation points in the signal that may correspond to defect locations. Taking the pile foundation with reduced diameter defects as an example, Fig. 5 shows the time cross-sectional changes of the reflected waves of such pile foundation defects and the typical waveform results after LPS processing.

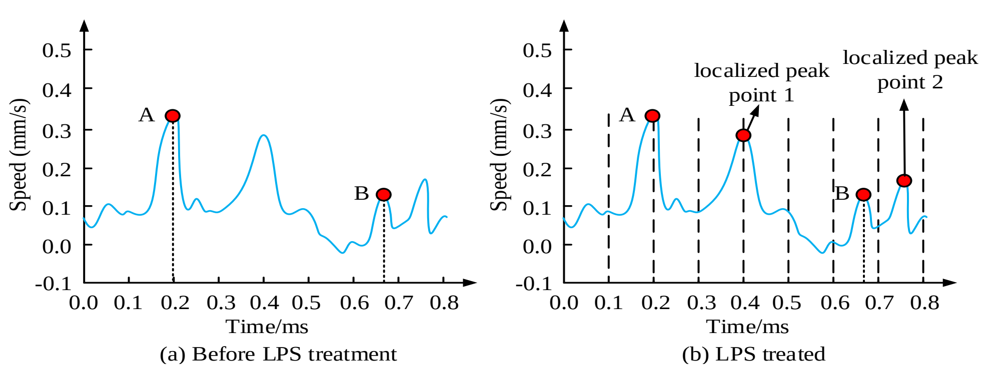



Fig. 5. Time-domain waveforms before and after LPS treatment

Figs. 5(a) and (b) show the time-domain waveforms of pile reflection before and after LPS treatment. In Fig. 5(a), point A corresponds to the initial wave of excitation-emission at the top of the pile, while point B is the location, where the wave reflects at the cross-section of the pile body, where it encounters a shrinkage defect. At this point, in addition to the main reflection peak, there are still multiple sets of interference waves with similar amplitudes in the signal, which makes the defect localization judgment uncertain. In Fig. 5(b), LPS forms a peak set by sliding judgment throughout the entire time series, comparing and extracting all peak points and their index positions that meet the conditions point by point. At this point, point A still represents the emission starting point, and the defect reflection peak corresponding to point B is more clearly marked. The interference peak is excluded or weakened because it does not meet the peak condition. This study introduces a second-order dynamic trend function to determine the rising or falling trend of the waveform at the current position, and defines a peak indicator function based on window weights, as shown in Eq. (10):

$$\psi(i) = \alpha \cdot \left( s_i - \frac{s_{i-\omega} + s_{i+\omega}}{2} \right) + \beta \cdot \left( \frac{d^2 s_i}{dt^2} \right) \tag{10}$$

In Eq. (10), $\psi(i)$ is the peak indicator function, and when $\psi(i) > 0$ is present, it is considered a candidate peak point. $\alpha$ and $\beta$ are both second-order differences of the signal. $\omega$ is the symmetrical window width. $s_i$ is the sub signal decomposed by WPT. To distinguish between real defect peaks and weak background disturbances, this study defines a weighted energy index $E_i^*$ and evaluates the sharpness of each candidate peak by combining signal gradient constraints, as shown in Eq. (11):

$$E_i^* = \frac{|s_i|^\gamma}{1 + \left| \frac{ds_i}{dt} \cdot \frac{d^2 s_i}{dt^2} \right|} \tag{11}$$

In Eq. (11), $\gamma$ is the amplitude control parameter. The points, where $E_i^*$ is greater than the threshold are ultimately retained as local reflection peaks for processing atypical defect reflection waveform such as shrinkage or mud inclusion in the pile body. This study combines WOA-LSTM and WPT-LPS to construct a novel PDI model, as shown in Fig. 6.

In Fig. 6, the entire PDI process is divided into two stages, corresponding to preliminary classification and fine recognition. Firstly, the model performs preprocessing operations such as normalization and denoising on the collected raw pile foundation detection signals to unify the data format and improve signal quality. Secondly, the preprocessed signal is input into an LSTM classification model optimized by WOA, which automatically extracts temporal dependent features and completes preliminary classification and discrimination between defects and non-defects. Subsequently, for the subset of signals judged as "defective" by WOA-LSTM, WPT multi-layer decomposition and LPS processing are performed sequentially. The full-frequency band energy characteristics and reflection peak position information are extracted to achieve further fine identification of defect types and structural features. In the end, the model outputs a comprehensive judgment result including the existence of defects, specific types, key reflection feature points, etc.

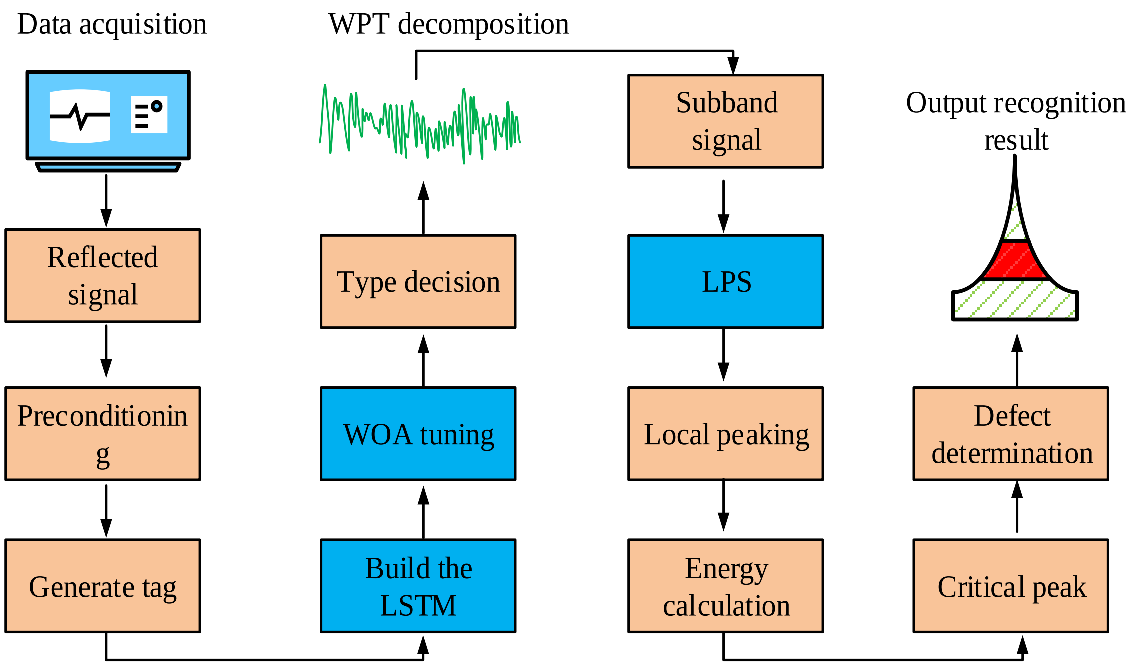
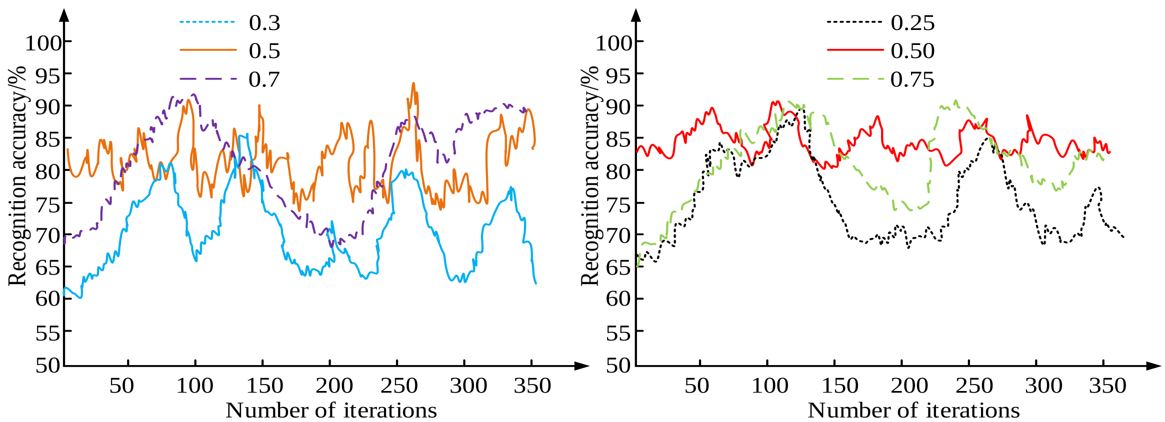



Fig. 6. New model flow of foundation PDI

## III. RESULTS

### A. Performance Testing of the New PDI Model

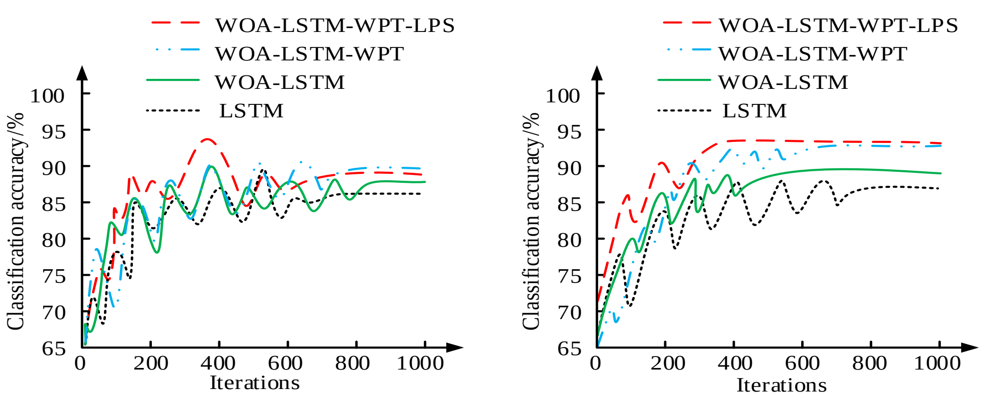
This study sets up a suitable experimental environment, with an Intel Core i7-12700H CPU, a clock speed of 2.3 GHz, a Windows 11 system, and 32 GB of memory. The GPU adopts NVIDIA RTX 3080 (16 GB of video memory) and the development environment is Python 3.10. The deep learning framework uses TensorFlow 2.12 and Keras 2.9. The Low Strain Pile Integrity Test Dataset (LSPIT) and Pile Sonic Logging Defect Imaging Dataset (PSLDID) are used as the testing data sources for pile foundation low strain integrity testing. Among them, LSPIT collects 1D time series signals of reflection waveforms of different types of foundation piles, such as intact piles, reduced diameter piles, broken piles, and mud-filled piles, under low-strain testing conditions. PSLDID mainly comes from the acoustic transmission method detection records in multiple large bridge and high-rise engineering projects at home and abroad. The data are in the form of 2D grayscale images, simulating the attenuation and abnormal distribution of sound waves on the propagation path inside the pile. This study first conducts value selection tests on the two types of hyperparameters that have the greatest impact on model performance, as shown in Fig. 7.




(a) Spiral contraction factor test (b) Amplitude control parameter testing

Fig. 7. Hyperparameter selection test result

Figs. 7(a) and (b) show the test results of selecting values for spiral contraction factor and amplitude control parameters. In Fig. 7(a), when the spiral contraction factor is set to 0.5, the overall model exhibits better convergence stability and recognition accuracy. Its accuracy rapidly improves in the early stages of iteration and remains at a high level of over 92.3% after 250 rounds. When the value is set to 0.3, although there are short-term high values in some sections, the overall fluctuation is large and the stability is poor. When the value is set to 0.7, the fluctuations in the first 200 rounds are relatively mild, but the final accuracy does not continue to improve, and the overall performance is slightly inferior to when the value is set to 0.5. In Fig. 7(b), when the amplitude control parameter is set to 0.50, the overall accuracy curve is relatively stable and remains above 90.8%. This indicates that the setting can balance global exploration and local convergence capabilities during the search process. Compared to others, when the value is set to 0.25, the model falls into early oscillations, with a large range of accuracy fluctuations and a tendency to fall into non-optimal regions. When the value is 0.75, the accuracy slightly improves in the middle and later stages, but overall it is not significantly better than 0.25. Therefore, based on the results of the two sets of tests, this study ultimately selects a spiral contraction factor of 0.5 and an amplitude control parameter of 0.5 as the recommended configurations for the WOA optimization module. Fig. 8 continues the ablation test.




(a) LSPIT (b) PSLDID

Fig. 8. Ablation test results

Figs. 8(a) and (b) show the ablation test values in the LSPIT and PSLDID datasets. In Fig. 8(a), WOA-LSTM-WPT-LPS consistently maintains the highest recognition accuracy and reaches a stable state around 750 rounds, with an accuracy rate of over 95.4%. In contrast, the WOA-LSTM-WPT model without an LPS module is slightly inadequate in high-frequency detail recognition, with an accuracy slightly lower by about 2 percentage points. In Fig. 8(b), the complete model exhibits fast convergence ability in the early stages and achieves an accuracy rate of over 96.7% after 700 rounds. After removing the LPS module, the structural expression ability of local reflection defects in the image decreases, and the model shows a slight lag. The comprehensive testing of two datasets shows that WPT and LPS modules have significant gain effects on defect time-frequency feature extraction and structural mutation recognition. The WOA optimization mechanism enhances the overall generalization ability and convergence stability of the model. Advanced models such as 3D-CNN, Empirical Mode Decomposition (EMD), and PDI Model Based on Apparent Wave Velocity Inverse Analysis (AWVIA-Pile) are introduced for comparison. Testing is conducted using precision (P), recall (R), F1 value, and average recognition time as indicators, as listed in Table I.

TABLE I. INDEX TEST RESULTS OF DIFFERENT MODELS

| Dataset | Model | P/% | R/% | F1 value/% | Average recognition time/s |
|---|---|---|---|---|---|
| LSPIT | 3D-CNN | 88.73 | 85.96 | 87.32 | 1.42 |
| | EMD | 84.29 | 80.67 | 82.44 | 1.87 |
| | AWVIA-Pile | 86.15 | 83.71 | 84.91 | 2.36 |
| | Research model | 94.62 | 92.85 | 93.73 | 0.97 |
| PSLDID | 3D-CNN | 89.54 | 86.43 | 87.96 | 1.57 |
| | EMD | 83.78 | 81.06 | 82.45 | 1.91 |
| | AWVIA-Pile | 85.41 | 83.28 | 84.33 | 2.14 |
| | Research model | 95.16 | 93.87 | 94.51 | 1.02 |

In Table I, on LSPIT, the P-value of the research model reaches 94.62%, the R-value is 92.85%, and the F1 value is as high as 93.73%, all significantly higher than the other three methods. In contrast, the EMD model has an F1 value of only 82.44% on this dataset, indicating poor recognition robustness under high noise interference. Although AWVIA-Pile has certain theoretical advantages, it has bottlenecks in practical recognition efficiency, with an average recognition time of 2.36s, significantly higher than the 0.97s of the research model. Similarly, on PSLDID, the research model still maintains a leading position, with an F1 value of 94.51% and a recognition time of 1.02 s that balances efficiency and accuracy. Although 3D-CNN has a certain spatial perception ability in image dimension modeling, its R-value is only 86.43% and its stability is slightly inferior. Therefore, the proposed model has good generalization ability and response efficiency while maintaining high-precision recognition.

### *B. New PDI Model Simulation Testing*

To verify the practical application effect of the model, this study simulates sand and clay foundation conditions and observes whether different models are affected by background signal interference on the decomposition ability of image temporal signal features, as shown in Fig. 9.

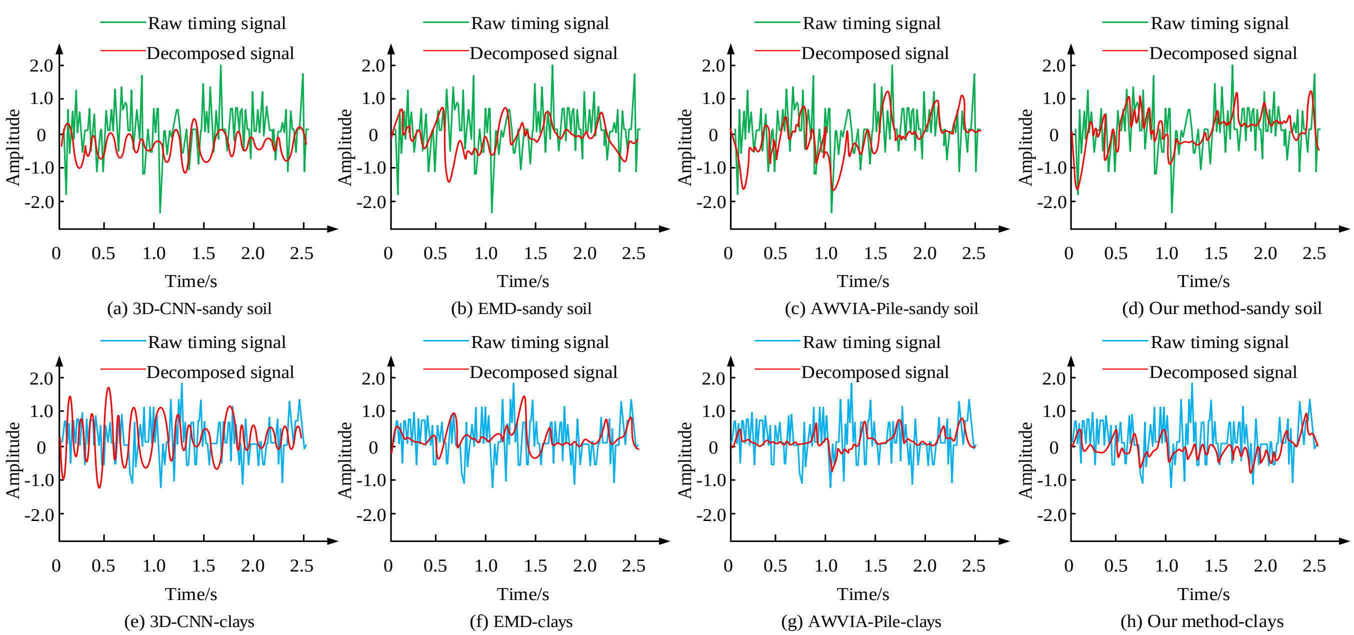
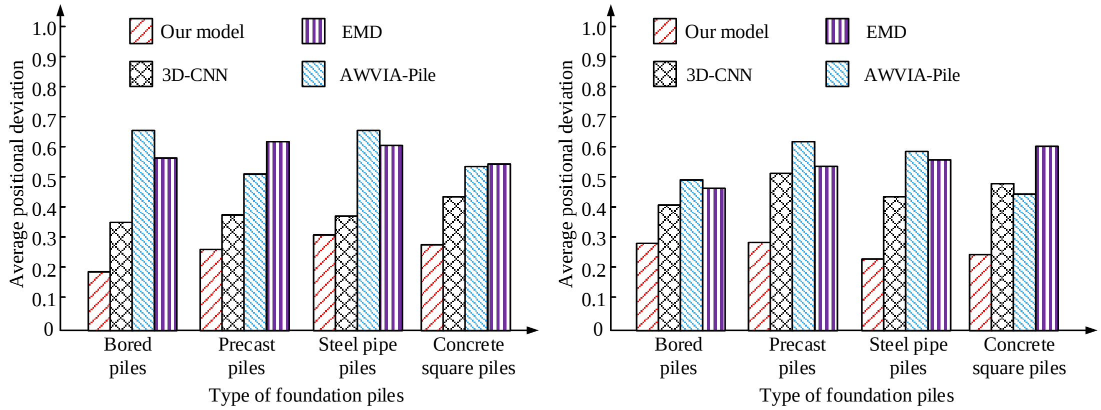



Fig. 9. Signal decomposition and comparison of each model in sandy soil and clay environment




(a) LSPIT (b) PSLDID

Fig. 10. Error results of defect identification of different pile foundations.

Figs. 9(a) to (d) and (e) to (h) show the signal analysis of four types of models in sandy and clay environments. In Figs. 9(a) to (d), in sandy soil environments, the 3D-CNN model only extracts low-frequency skeleton structures and lacks the ability to respond to high-frequency abrupt signals. Although EMD has a certain deconstructive ability, its decomposition results exhibit a signal drift phenomenon. The feature curves extracted by AWVIA-Pile exhibit weak points such as edge blurring and energy collapse, showing sensitivity to shallow noise. The research method shows a clearer and more clearly defined signal decomposition effect, which not only preserves the complete structure of the reflected main wave but also effectively weakens the interference of background noise, indicating that the model has stronger time-frequency separation ability in interference environments. In Figs. 9(e) to (h), the models are all affected by more complex background waveforms in clay environments, resulting in a significant increase in decomposition difficulty. However, the research model still maintains a high decomposition resolution, with clear hierarchical structures of the main and secondary waves, and prominent defect band characteristics. This indicates that it also has good structural preservation ability and noise adaptability in low-permeability formations. This study takes drilled pile, precast pile, steel pipe pile, and concrete square pile as examples to test the average position deviation of defect detection for each model, as shown in Fig. 10.

Fig. 10(a) shows the pile foundation defect identification errors of four methods in the LSPIT and PSLDID datasets. In Fig. 10(a), in the identification task of the drilled pile, the average position deviation of the research model is the smallest, only 0.21. Compared with 3D-CNN, EMD, and AWVIA Pile methods, it reduces by about 0.15, 0.22, and 0.18, indicating stronger feature locking ability under complex multi-wave interference conditions. Due to the high regularity of signal reflection patterns in concrete square piles, the overall error of each model is slightly smaller, but the research model still outperforms the comparison method with a minimum deviation of 0.24. In Fig. 10(b), the positional deviation of the research model on four types of pile types is controlled below 0.3, with errors of 0.19 and 0.23 for drilled pile and precast pile, which are much lower than the fluctuation range of 3D-CNN and AWVIA Pile models. This indicates that it has stronger localization robustness in defect area determination of image data. This study tests single and multiple defects based on recognition accuracy, model complexity, and average recognition delay, as shown in Table II.

TABLE II. SINGLE DEFECT AND MULTIPLE DEFECT INDEX TEST RESULTS

| Number of defects | Model | Precision/% | Model complexity/% | Average recognition time/s |
|---|---|---|---|---|
| Single defect | 3D-CNN | 91.47 | 82.53 | 0.76 |
| | EMD | 88.92 | 69.41 | 0.89 |
| | AWVIA-Pile | 90.26 | 76.89 | 0.93 |
| | Research model | 96.38 | 65.28 | 0.58 |
| Multiple defects | 3D-CNN | 86.51 | 82.53 | 0.81 |
| | EMD | 83.64 | 69.41 | 0.94 |
| | AWVIA-Pile | 85.23 | 76.89 | 0.97 |
| | Research model | 93.27 | 65.28 | 0.63 |

In Table II, when faced with a single defect recognition task, the research model achieves a recognition precision of 96.38%, which is 4.91% and 6.12% higher than 3D-CNN and AWVIA-Pile. Meanwhile, the model complexity is only 65.28%, indicating that the network structure is lighter while maintaining recognition ability, and the average recognition delay is 1.14 s, significantly better than AWVIA-Pile's 2.03 s. In the multi-defect recognition task, the research model still maintains a high recognition precision of 93.27%, while 3D-CNN and EMD show significant fluctuations under multi-target interference, with precision decreasing to 86.51% and 83.64%, and recognition time both exceeding 1.4 s. In addition, the delay of AWVIA-Pile increases to 2.19s in multi-defect scenarios, indicating its weak structural decoupling ability for composite defects. In summary, this research method has good stability and precision control ability in single defect scenarios, and exhibits stronger adaptability and efficiency advantages in complex tasks with multiple defects, with high practical application value and promotion prospects.

## IV. Conclusion

In response to the problems of insufficient classification accuracy, weak noise resistance, and inaccurate structural localization in the current PDI process, this study constructed a defect data classification method by combining LSTM and WOA. At the same time, a novel PDI model was proposed by combining WPT and LPS for temporal feature decomposition and recognition of defect labels. In the experiment, when both the spiral contraction factor and amplitude control parameters were set to 0.5, the recognition accuracy of the model remained at a maximum of 92.3%. Compared with simple LSTM and WPT, after sequentially combining WOA and LPS, the final combined model achieved the highest classification accuracy of 96.7%, showing a significant improvement effect. Compared to other models, this new method achieved the highest P, R, and F1 values of 95.16%, 93.87%, and 94.51%, and the shortest average recognition time of 0.97s. Under sandy and clay foundation conditions, the signal decomposition effectiveness of the research method was higher, and the decomposed sub-signals were clearer and more realistic. For the four typical types of PDI, the accuracy was higher, especially for drilled and precast piles with errors of 0.19 and 0.23, which were much lower than other methods. The lowest complexity could reach 65.28%, and the shortest average recognition delay was 0.58s, both demonstrating excellent processing efficiency and effectiveness. In summary, the new method performs particularly well in handling data types with relatively regular structures and obvious signal characteristics such as drilled piles and prefabricated piles, with better positioning errors and recognition accuracy than other types of piles. For data with higher signal complexity or more diverse defect types, the algorithm is still well adapted. However, the study still has some limitations. First, the generalization ability of the model needs to be further improved when facing extreme working conditions and unseen defect types. Second, current recognition is mainly based on single modal signals, and in the future, multimodal fusion can be considered to enhance the robustness and adaptability of the model. In addition, the integration and optimization of the model with the actual inspection equipment needs to be enhanced to improve the convenience and real-time performance of engineering applications. Future research will consider introducing multimodal data augmentation mechanisms, transfer learning strategies, and integrated optimization with actual detection devices to further promote the application of this model in actual bridge and building pile foundation detection.

## References

[1] H. Shen, X. Li, R. Duan, Y. Zhao, J. Zhao, H. Che, et al., “Quality evaluation of ground improvement by deep cement mixing piles via ground-penetrating radar.” Nature Communications, vol. 14, no. 1, pp. 3448-3452, 2023.

[2] J. Wang, H. Zhu, D. Tan, Z. Li, C. Wei, and B. Shi, “Thermal integrity profiling of cast-in-situ piles in sand using fiber-optic distributed temperature sensing.” Journal of Rock Mechanics and Geotechnical Engineering, vol. 15, no. 12, pp. 3244-3255, 2023.

[3] J. Wu, M. H. El Naggar, and K. Wang, “Pile damage detection using machine learning with the multipoint traveling wave decomposition method.” Sensors,vol. 23. no. 19, pp. 8308-8312, 2023.

[4] W. Zhang, K. Zhu, Z. Yang, J. Ding, and J. Gan, “Development of an underwater detection robot for the structures with pile foundation.” Journal of Marine Science and Engineering, vol. 12, no. 7, pp. 1051-1058, 2024.

[5] S. Jiang, W. Wang, Z. Su, and S. Wang, “Automatic detection of surface defects on underwater pile - pier of bridges based on image fusion and deep learning.” Structural Control and Health Monitoring, vol. 6,no. 1, pp. 84290-84293, 2023.

[6] H. Liu, W. Wu, X. Yang, X. Liu, L. Wang, M. H. E. Naggar, et al., “Apparent wave velocity inverse analysis method and its application in dynamic pile testing.” International Journal for Numerical and Analytical Methods in Geomechanics, vol 47, no. 4, pp. 549-569, 2023.

[7] X. Wang, Z. Qin, X. Bai, Z. Hao, N. Yan, andJ. Han, “Research progress of machine learning in deep foundation pit deformation prediction.” Buildings, vol. 15, no. 6,pp. 852-859, 2025.

[8] C. S. Wu, M. Ge, L. L. Qi, D. Zhuo, J. Zhang, T. Hao, et al., “Multi-defect identification of concrete piles based on low strain integrity test and two-channel convolutional neural network,” Appl. Sci., vol. 13, no. 6, pp. 3530–3537, 2023.

[9] H. Wang, S. Zhang, J. Li, Y. Yuan, and F. Zhang, “Classification of low-strain foundation pile testing signal using recurrent neural network,” Buildings, vol. 13, no. 5, pp. 1228–1291, 2023.

[10] J. Wu, M. H. El Naggar, and K. Wang, “A hybrid convolutional and recurrent neural network for multi-sensor pile damage detection with time series,” Sensors, vol. 24, no. 4, pp. 1190–1194, 2024.

[11] T. Hu and J. Xu, “Prediction of buildings’ settlements induced by deep foundation pit construction based on LSTM-RA-ANN,” Appl. Sci., vol. 14, no. 12, pp. 5021–5033, 2024.

[12] Shen S, Zeng Y, Lai C. Rapid three-dimensional reconstruction of underwater defective pile based on two-dimensional images obtained using mechanically scanned imaging sonar. Structural Control and Health Monitoring, 2023, 2023(1): 36474-36479.

[13] Y. Wu, F. Xiao, F. Liu, Y. Sun, X. Deng, L. Lin, et al., “A visual fault detection algorithm of substation equipment based on improved YOLOv5,” Appl. Sci., vol. 13, no. 21, pp. 11785–11788, 2023.

[14] Y. Cao, J. Ni, J. Chen, and Y. Geng, “Rapid evaluation method to vertical bearing capacity of pile group foundation based on machine learning,” Sensors, vol. 25, no. 4, pp. 1214–1217, 2025.

[15] A. Picardo, M. A. Millán, R. Galindo, and A. Alencar, “Revisiting the analytical solutions for ultimate bearing capacity of pile embedded in rocks,” J. Rock Mech. Geotech. Eng., vol. 15, no. 6, pp. 1506–1519, 2023.

[16] G. Chen, Y. Wang, X. Li, Q. Bi, and X. Li, “Shovel point optimization for unmanned loader based on pile reconstruction,” Comput.-Aided Civ. Infrastruct. Eng., vol. 39, no. 14, pp. 2187–2203, 2024.

[17] D. Chen, J. Zhou, P. Duan, and J. Zhang, “Integrating knowledge management and BIM for safety risk identification of deep foundation pit construction,” Eng. Constr. Archit. Manage., vol. 30, no. 8, pp. 3242–3258, 2023.

[18] D. Cardenas, P. Loncomilla, F. Inostroza, P. Tsunekawa, and J. Ruiz-del-Solar, “Autonomous detection and loading of ore piles with load-haul-dump machines in Room & Pillar mines,” J. Field Robot., vol. 40, no. 6, pp. 1424–1443, 2023.

[19] N. Li, T. Ye, Z. Zhou, C. Gao, and P. Zhang, “Enhanced YOLOv8 with BiFPN-SimAM for precise defect detection in miniature capacitors,” Appl. Sci., vol. 14, no. 1, pp. 429–431, 2024.

[20] S. Pal, A. Roy, P. Shivakumara, and U. Pal, “Adapting a swin transformer for license plate number and text detection in drone images,” Artif. Intell. Appl., vol. 1, no. 3, pp. 145–154, 2023.