Automatic Quantification of Subsurface Defects by Analyzing Laser Ultrasonic Signals Using Convolutional Neural Networks and Wavelet Transform

Shifeng Guo, Haowen Feng, Wei Feng, Gaolong Lv, Dan Chen, Yanjun Liu, and Xinyu Wu, *Member, IEEE*

***Abstract*— The conventional machine learning algorithm for analyzing ultrasonic signals to detect structural defects necessarily identifies and extracts either time- or frequency-domain features manually, which has problems in reliability and effectiveness. This work proposes a novel approach by combining convolutional neural networks (CNNs) and wavelet transform to analyze the laser-generated ultrasonic signals for detecting the width of subsurface defects accurately. The novelty of this work is to convert the laser ultrasonic signals into the scalograms (images) via wavelet transform, which are subsequently utilized as the image input for the pretrained CNN to extract the defect features automatically to quantify the width of defects, avoiding the necessity and inaccuracy induced by artificial feature selection. The experimentally validated numerical model that simulates the interaction of laser-generated ultrasonic waves with subsurface defects is first established, which is further utilized to generate adequate laser ultrasonic signals for training the CNN model. A total number of 3104 data are obtained from simulation and experiments, with 2480 simulated signals for training the CNN model and the remaining 620 simulated data together with 4 experimental signals for verifying the performance of the proposed algorithm. This approach achieves the prediction accuracy of 98.5% on validation set, particularly with the prediction accuracy of 100% for the four experimental data. This work proves the feasibility and reliability of the proposed method for quantifying the width of subsurface defects and can be further expanded as a universal approach to various other defects detection, such as defect locations and shapes.**

***Index Terms*— Convolutional neural networks, laser ultrasonic signal, nondestructive evaluation (NDE), numerical model, subsurface defects, wavelet transform.**

Manuscript received May 8, 2021; accepted June 6, 2021. Date of publication June 9, 2021; date of current version September 27, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 52071332, in part by the Joint Research Fund between the National Natural Science Foundation of China and Shenzhen under Grant U1813222, in part by the Joint Research Fund between the National Natural Science Foundation of China and Hebei Province under Grant U20A20283, in part by the Department of Science and Technology of Guangdong Province under Grant 2019QN01H430 and Grant 2019TQ05Z654, in part by the National Key Research and Development Program of China under Grant 2018YFB1309403, in part by the Science and Technology Innovation Commission of Shenzhen under Grant ZDSYS20190902093209795 and Grant JCYJ20180507182239617, and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2020A1515110218. *(Corresponding authors: Wei Feng; Xinyu Wu.)*

Shifeng Guo is with the Shenzhen Key Laboratory of Smart Sensing and Intelligent Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China, also with the University of Chinese Academy of Sciences, Beijing 100049, China, also with the Guangdong Provincial Key Lab of Robotics and Intelligent System, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China, and also with the CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: sf.guo@siat.ac.cn).

Haowen Feng is with the Shenzhen Key Laboratory of Smart Sensing and Intelligent Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: hw.feng@siat.ac.cn).

Wei Feng is with the University of Chinese Academy of Sciences, Beijing 100049, China, and also with the CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: wei.feng@siat.ac.cn).

Gaolong Lv and Dan Chen are with the the Shenzhen Key Laboratory of Smart Sensing and Intelligent Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China, and also with the Guangdong Provincial Key Lab of Robotics and Intelligent System, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: gaolonglv@siat.ac.cn; danchen@siat.ac.cn).

Yanjun Liu is with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: yjliu@sustc.edu.cn).

Xinyu Wu is with the Guangdong Provincial Key Laboratory of Robotics and Intelligent System, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China, and also with the CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Shenzhen 518055, China (e-mail: xy.wu@siat.ac.cn).

Digital Object Identifier 10.1109/TUFFC.2021.3087949

I. Introduction

THE subsurface defects hidden in metallic structures are detrimental to the integrity of critical structures and may cause ruinous damage ultimately [1], [2]. A reliable and effective nondestructive evaluation (NDE) technique to detect and size the subsurface defects is therefore crucial for securing safety and operation. Ultrasound has been widely implemented in NDE practice for obtaining information, including the location and size of subsurface defects in different materials [3], [4]. However, the conventional ultrasonic inspection is generally contact-based and requires additional fixtures or coupling for ultrasonic signal generation and acquisition [5], which is therefore not applicable in a harsh environment, such as high temperature and heavy radiation [5].

The laser ultrasonic technique, using a pulsed laser to generate and adopting a continuous laser interferometer to detect the ultrasonic signals, is a pure noncontact detection technique and has been widely utilized for inspecting structural materials with various defects. Lévesque *et al.* [6] used laser ultrasonics combining with the synthetic aperture focusing technique to analyze the additive manufactured coupons. Bate *et al.* [7] implemented laser ultrasonics to obtain information of microstructures of metallic alloy. Selim *et al.* [8] proposed a hybrid system combining remotely induced laser ultrasonic with conventional transducer for defects detection in metallic parts. Zeng *et al.* [9] testified the feasibility of using laser ultrasonics to inspect additive manufactured components. The abovementioned work generally relies on establishing a linear correlation between the detected defects and the parameters of the received ultrasonic signals to obtain the defect information. However, in practical applications, it is extremely expensive and time-consuming to obtain the correlation between defects and ultrasonic signals by a mass of experiments and samples.

The approaches combining machine learning with the NDE technique that enables establishing the complex nonlinear relationship between defects and ultrasonic parameters via neural networks are emerging in recent years [10], [11]. Zhang *et al.* [2] quantified the subsurface defects by combining an improved genetic algorithm–backpropagation neural network (GA-BPNN) with laser ultrasonic technique. Krummenacher *et al.* [12] proposed two machine learning methods to automatically detect the wheel cracks by a pretraining artificial neural network (ANN) with convolutional layers on 2-D representations of the measurement time series for learning features and using support vector machines (SVMs) for classification. Li *et al.* [1] used the features of laser-generated surface acoustic waves as inputs to a neural network optimized with particle swarm optimization (PSO) algorithm to predict the depth of surface defects. All the above works necessarily require to select and extract either time- or frequency-domain features from detected ultrasonic signals for training the machine learning algorithm. However, this procedure is very subjective and empirical, and in most cases, it is very difficult to extract the reliable features that are most effective or relevant with the defects to be detected.

A robust method that can automatically identify the most relevant features of ultrasonic signal with defects and eliminate the subjective selection is therefore highly demanded. Deep learning, a new area in machine learning, which can extract and identify the effective features of images automatically, may be a potential method to address the above issue. In deep learning, convolutional neural networks (CNNs) have proved to have excellent performance for visual recognition tasks and are widely implemented [13]. Makantasis *et al.* [14] used a deep learning technique to recognize the human activities in industrial environments by appropriately transforming the video input into incorporate temporal information into each frame. Maninis *et al.* [15] constructed a deep CNNs architecture for both retinal vessel and optic disk segmentation through studying the eye fundus images. However, CNNs require high-resolution images as input, which imposes challenges for conventional ultrasonic NDE technique with only time-domain ultrasonic signals obtained. The wavelet transform technique enables the transformation of time-domain signals into images (scalogram) for CNN analysis. The scalogram is the time–frequency representation of the time-domain signal through wavelet transform, and the coefficient values at the corresponding time–frequency locations can be revealed by its color and brightness. The scalograms (images) with CNNs analysis has been successfully implemented in research on various kinds of time-domain signals, such as acoustic signals [16]–[18] and electroencephalogram (EGG) [19]. Ren *et al.* [18] transformed the acoustic signal into scalograms based on the morse and bump wavelet bases and used the scalograms as the input of pretrained CNNs and (bidirectional) gated recurrent neural network to classify the acoustic scene. The application of CNN based on time–frequency images, either obtained from piezoelectric or pulsed laser ultrasonic waves, for detecting structural defects has not been reported yet. As the time-domain laser ultrasonic signals are similar to the above acoustic signals, combining wavelet transform and CNN can serve as a new effective way to analyze the ultrasonic signals to detect structural defects.

In this work, a systematic approach that combines CNN and wavelet transform is proposed to evaluate subsurface defects in metallic alloy materials by analyzing laser-generated ultrasonic signals. Unlike conventional machine learning-based work on defects detection necessarily extracting either time- or frequency-domain ultrasonic signals as features manually, it is very subjective and unreliable in many cases. This method enables the quantification of the width of subsurface defects automatically, avoiding the necessity and inaccuracy introduced by artificial feature selection. The novelty of this work is to convert the laser ultrasonic signals into the scalograms (images) via wavelet transform and the obtained scalograms are subsequently utilized as the image input for the pretrained CNN to quantify the width of defects. Considering that CNN requires large training data to assure the accuracy of prediction, it is extremely expensive and time-consuming to obtain such a large number of data by experiments. Therefore, a reliable numerical model is established to obtain sufficient simulation data to prove the concept of this work. The numerical model is first established and experimentally validated to simulate the interaction of laser-generated ultrasonic waves with subsurface defects. The obtained ultrasonic signals are thereafter converted into the scalograms via reliable wavelet transform. Three pretrained CNNs from large-scale datasets ImageNet are further introduced to extract the deep features of scalograms from ultrasonic signals with characteristics of subsurface defects automatically. Finally, the multiplicative long short-term memory network (mLSTM) is chosen as the classifier for predicting the crack widths, by receiving the features extracted using the above pretrained CNNs. The framework of this work is shown in Fig. 1.

II. Materials and Method

A. Specimens and Experiment Data

The aluminum alloy (2024) plates with dimension of 30 mm in length, 10 mm in width, and 5 mm in height were used

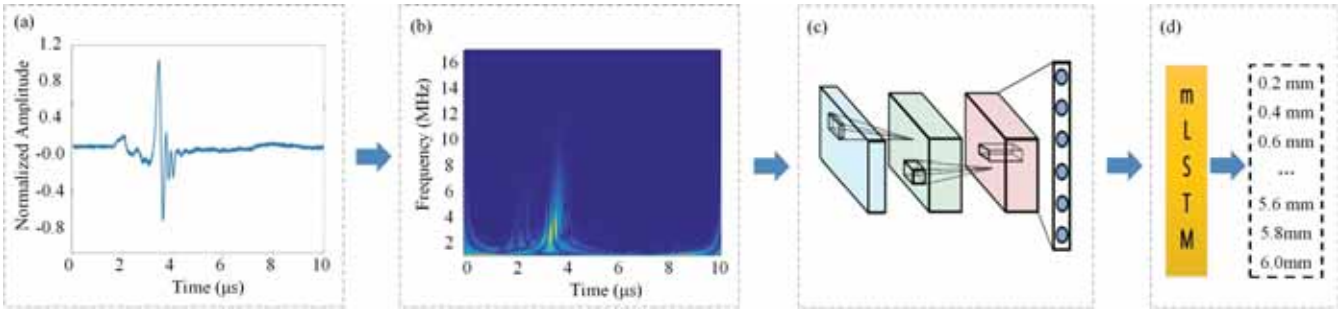

Fig. 1. Flowchart of the combination of the CNNs and laser ultrasonic technique for detecting the width of subsurface cracks. (a) Simulated time-domain laser ultrasonic signal with defect width of 4 mm. (b) Scalogram obtained from wavelet transform of the laser ultrasonic signal. (c) Image features extraction by pretrained CNNs. (d) mLSTM for defect width prediction.

for the experiments. The subsurface defects with fixed defect height (H) of 1 mm, depth-to-surface (D) of 1 mm, and various defect widths (W) of 0, 2, 3, and 4 mm were fabricated using electrical discharge machining and wire cutting method. The schematic of the laser ultrasonic experimental setup is shown in Fig. 2(a). A Nd:YAG pulsed laser (Centurion+) with the wavelength of 1064 nm, pulse duration of 12 ns, and pulse energy of 5 mJ is utilized to generate Rayleigh ultrasonic waves, and a laser interferometer (TEMPO 2D, Sound and Bright, USA) is adopted to detect the ultrasonic waves. The pulsed laser is excited at 5 mm away from the center of subsurface defects, and the detector is positioned at 10 mm from it to receive the transmitted Rayleigh ultrasonic waves. The acquired signals are digitized with a sampling frequency of 250 MHz and averaged over 300 times to improve the signal-to-noise ratio. The laser-generated ultrasonic signals are substantially affected by the width of defects, both in amplitude and phase, as shown in Fig. 2(b). To minimize the effect of surface roughness, the surface of all experimental specimens is maintained at about 0.8 μm.

B. Numeric Model and Simulation Data

The simulation is conducted using the thermal stress module with transient solver in Comsol Multiphysics 5.4. All materials are assumed to be isotropic linear elastic materials, and no damping is considered. According to the shape of laser source and the shape of defect, the model in our work is simplified as a 2-D axisymmetric model. In order to simulate the laser radiation effect, a heat flux load is applied on the top side of the 2-D model. The artificial subsurface defect is modeled as a rectangle notch with a fixed height and depth of 1 mm and various widths varied from 0 to 6 mm with an increment of 0.2 mm, with a total of 31 simulation data obtained. The distance between the generation and detection laser is fixed at 10 mm. The parameters of the laser pulse spot radius, power density, and rise time are set to 1 mm, $5mW/cm^2$, and 12 ns, respectively. The material properties of 2024 aluminum alloy used in the numerical calculation are listed in Table I.

As the convergence of finite-element calculation is significantly affected by the temporal and spatial resolution of the numerical model [1], the appropriate mesh size (L_e) and proper time step (Δt) are the two important aspects to highly decide the fidelity and accuracy of the numerical simulation,

TABLE I
RELEVANT PARAMETERS OF 2024 ALUMINUM ALLOY MATERIALS FOR NUMERICAL SIMULATION

Parameters	Value
Density ρ/(kg/m^3)	2780
Young's Modulus E/(GPa)	71
Poisson's ratio υ	0.33
Thermal expansion α/(μK)	23
Thermal conductivity k/(Wk/m)	120
Heat capacity C_v/(JK/kg)	880
Reflection coefficient	0.9

which can be modeled as [20]

$$L_e = \frac{\lambda_{\min}}{10} \tag{1}$$

$$\Delta t = \frac{1}{20 f_{\max}} \tag{2}$$

where $f_{\max}$ and $\lambda_{\min}$ are the highest frequency and the shortest wavelength of the laser-generated ultrasonic waves, respectively. In the simulation model, the normal node displacement at the detection location is extracted for the acquisition of ultrasonic signal. The quadrilateral elements with four nodes are applied, and the mesh size near the laser irradiation area and the other mesh size away from the heated-affected area are arranged to be 2 and 20 μm, respectively, which can guarantee the continuous spread of ultrasound and satisfy the change of the temperature gradient. The time step and simulation time are set to 4 ns and 10 μs, respectively, which can ensure the accuracy of the numerical solution and meanwhile enables the representation of all phenomena implied in the simulation process. In addition, the boundary condition of both sides used in the numerical model is set as low-reflecting boundary to suppress the reflection of ultrasonic waves.

C. Wavelet Transform

To extract the time–frequency information hidden in the defect ultrasonic signal automatically, the scalogram of the signal through wavelet transform is utilized. Compared with the conventional time–frequency methods such as short-time Fourier transform (STFT), continuous wavelet transform (CWT) is a time-scale representation. In the scalogram, the wavelet coefficient values at the corresponding time–frequency locations are represented by the gray value

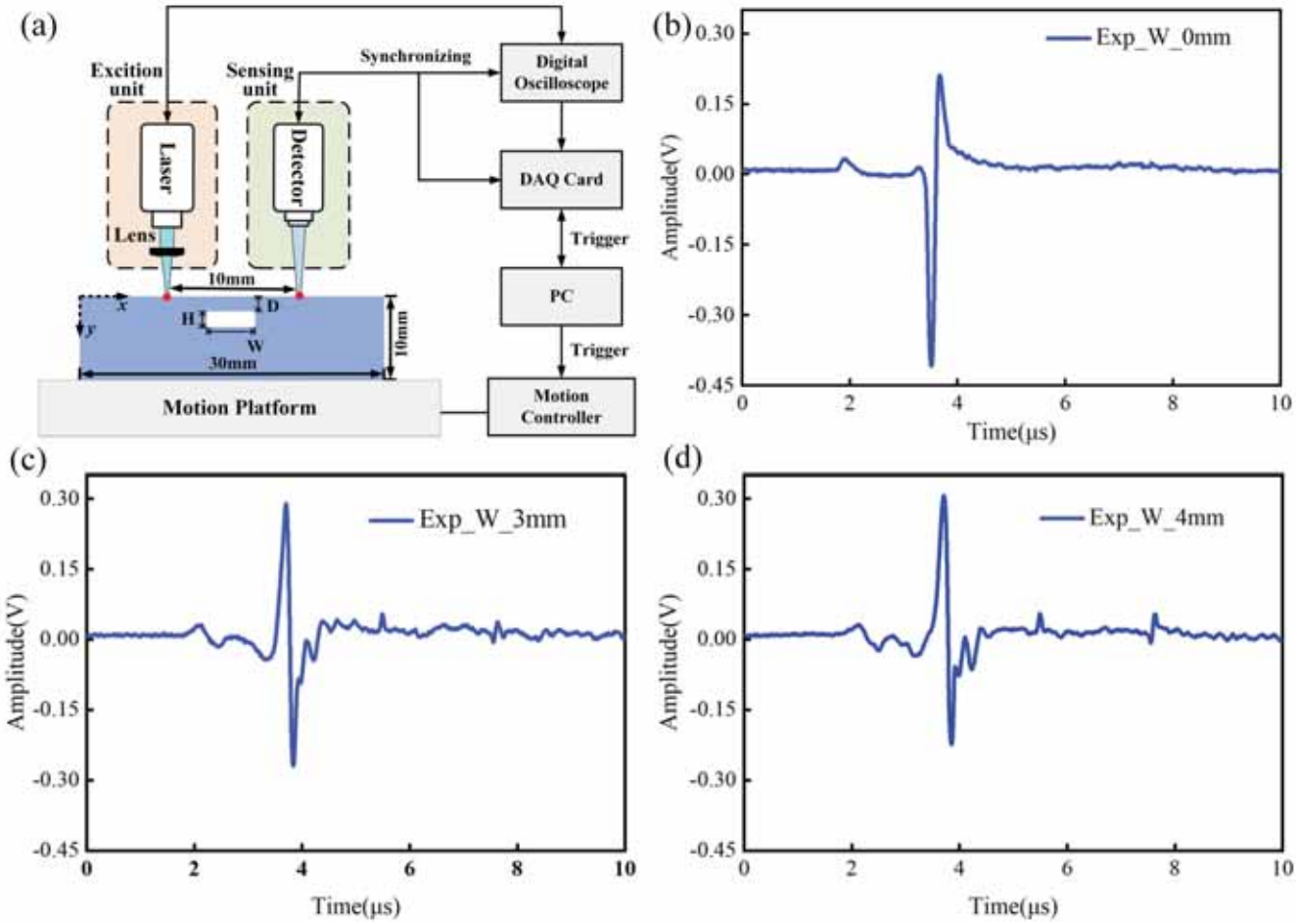

Fig. 2. (a) Schematic of laser ultrasonic detection of subsurface defects in metallic material. H, D, and W here are abbreviations of defect height, depth-to-surface, and defect width, respectively. (b)–(d) Representative laser ultrasonic signals obtained from specimens with various defect widths.

in the 2-D image, which exposes more detailed information of the signal.

For a given signal x(t), the CWT is defined as [21]

$$\mathrm{CWT}_x(\beta,\varepsilon)=\int_{-\infty}^{+\infty}x(t)\varphi^*_{\beta,\varepsilon}(t)dt \tag{3}$$

where $\beta \in R^+$ stands for the scale parameter, $\varepsilon \in R^+$ is the translation diameter of time shifting and the basis function, and $\varphi^*_{\beta,\varepsilon}$ is obtained by scaling the mother wavelet $\varphi(t)$ at time ε and scale β. The asterisk represents that the complex conjugate of the wavelet function is used in the transform. To select the optimal wavelet basis function for the laser-generated ultrasonic signals, five conventionally used wavelet basis functions that are most similar to the waveform of the defect signals, ***db4***, ***db10*** [22], ***coif3*** [23], ***bump*** [24], and ***morse*** [25] are utilized and compared. In this work, the wavelet transform of the signal is conducted by using the wavelet transform toolbox of MATLAB [26].

D. Deep Learning

1) Features Extraction With Pretrained CNNs: Due to the power of transfer learning technique (i.e., saving training time and improving the generalization capability of the network), the fine-tuning pretrained network has advantages over training a network with randomly initialized parameter weights [27]. Since using multiple convolution networks can extract image features more accurately and completely, "VGG16," "Resnet50," and "DenseNet161" are chosen as the pretrained CNNs for extracting the deep features of the scalogram obtained from wavelet transform after comparing the training results of different variants of these three models. These three deep CNNs have shown excellent performance in an enormous amount of image classification tasks, such as ImageNet Large Scale Visual Recognition Challenge (ILSVRC). The three pretrained CNNs were obtained from Pytorch's model library and were replaced by the last layer of them with a fully connected layer for concatenating. Then, the output parameters of the fully connected layer will be used as the input to mLSTM classifier with a softmax layer for defect width prediction.

The softmax activation function is defined as [28]

$$\mathrm{Softmax}(z_i)=\frac{e^{z_i}}{\sum_{k=1}^{K}e^{z_k}} \tag{4}$$

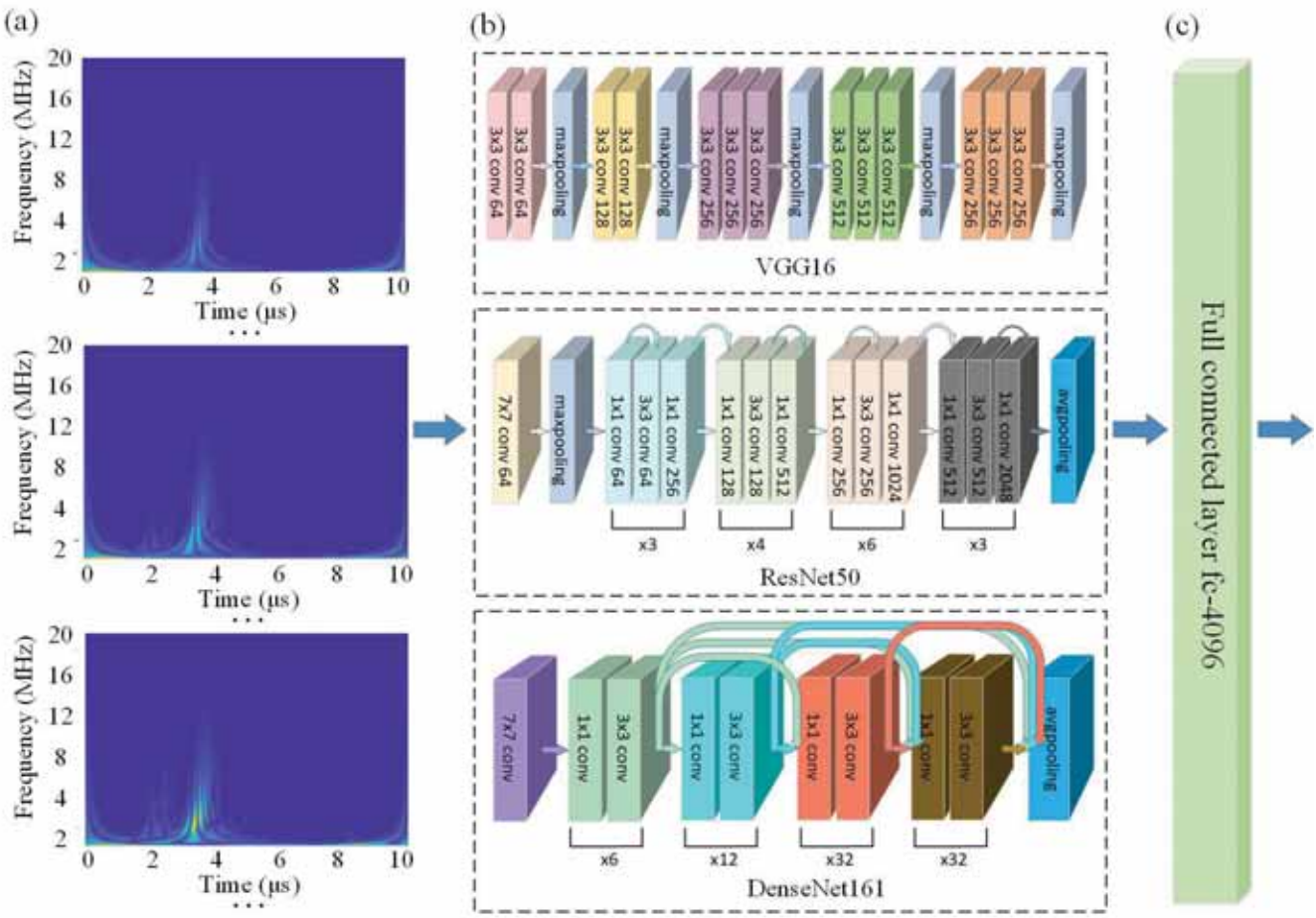

Fig. 3. (a) Scalograms obtained from wavelet transform. (b) Structure of three pretrained CNNs for feature extraction. (c) Fully connected layer to concatenate and output the extracted features to mLSTM for defect width prediction.

where z_i represents the output of the ith classification and K stands for the number of classification.

As shown in Fig. 3, the scalograms of laser ultrasonic signals with various defect widths obtained from wavelet transform are used as the input of three pretrained CNNs for deep feature extraction, and a fully connected layer is added to concatenate the extracted features and output them to mLSTM for defect width prediction automatically.

a) VGG16 model: Visual geometry group (VGG) network is a CNN proposed by the Oxford Visual Geometry Group, Oxford, U.K., which has excellence on image localization and classification [29]. In this network, several sizes of 3×3 convolution kernels are used to replace the larger convolution kernels (three 3×3 replace a 7×7 and two 3×3 replace a 5×5), which is shown in Fig. 3(b). In this way, the number of parameters between each network layer can be reduced substantially. Fewer parameters mean less overfitting, and more importantly, the stack of two 3×3 convolutions has more nonlinear transforms than a single 5×5 convolution, which enables the network to have an improved capability in feature learning. In this article, VGG16, a kind of VGG model that contains 16 convolution layers and has great learning performance in VGG series, is selected as one of the pretrained CNNs.

b) ResNet50 model: ResNet is the abbreviation of residual network, a network structure proposed by He *et al.* [30]. Since deep convolutional neural networks have made a series of breakthroughs for image classification and recognition, adding the depth of network has become a trend in the field of deep learning. However, notorious problems occur when training a deeper neural network is vanishing gradient problem and degradation problem. Various methods can be used to solve the vanishing gradient problem, including normalized initialization and intermediate normalization layers. However, the degradation problem remains a challenge to be solved. Residual learning is proposed to solve the degradation problem due to the merits of its special parameter delivery way. Instead of learning features, this model tries to learn some residuals. The residual learning structure is a shortcut connection that is shown in Fig. 3(b). Also, ResNet50 is a variant of ResNet, which is a 50-layer residual network [31].

c) DenseNet161 model: DenseNet is a dense convolutional network, which makes improvement on the basis of ResNet [32]. ResNet creates short paths from early layers to later layers to address the problem that gradient vanishes when passing many layers. However, the identification function and the output of each layer in ResNet are combined via adding, which may block the flow of information. To improve the prob-

lem of information flow between layers, DenseNet introduces the dense block and transition layer, as shown in Fig. 3(b). Each layer in the dense block can learn the features produced by the previous layers in forwarding propagation, which can alleviate the risk of exploding and vanishing gradient problem and accelerate the speed of gradient propagation. Due to this dense connectivity, the dimension of feature maps received by the subsequent layers increases largely as the depth of dense block grows. Therefore, there are transition layers set between two dense blocks for reducing the dimension with a 1×1 convolution kernel. In this article, DenseNet161, a species of DenseNet with 161 layers incorporated with other models exhibits excellence on features extraction [33].

2) mLSTM Classifier for Prediction: As a variant of LSTM, mLSTM is an amalgam structure that integrates the flexible input-independent transitions of mRNNs with the gating structure in LSTM [34]. Due to the combination architecture of mRNNs and LSTM, the convoluted transition produced by the factorized hidden weight matrix becomes much easier to control through the gating unit. Besides, more flexible input-independent transition functions could be applied since the additional sigmoid input and forget gates setting in LSTM. Since the distributed hidden features could be changed fast in mLSTM architecture, it surpasses other variants of LSTM and all previous networks in Hutter Prize and some language modeling tasks [35]. Therefore, mLSTM is selected as the classifier, which receives the scalogram features extracted by the above pretrained CNNs and predicts the corresponding defect width.

III. Result and Discussion

A. Experimental Results and Numerical Model Validation

A reliable and validated numerical model is crucial for obtaining adequate data for training the deep learning model. The dataset of this work is obtained from the simulation model of laser-generated ultrasonic waves with distinct subsurface defects (crack widths ranging from 0 to 6 mm at 0.2-mm interval, with 31 groups in total). The experimental and simulation data (from specimen with defect width of 2 and 3 mm) are, respectively, normalized in the range from -1 to 1 for comparison [36], as shown in Fig. 4. The results show that the simulated laser ultrasonic waves are in good consistent with the experimental signals, as highlighted in the dashed area. It therefore indicates that the established simulation model is validated and reliable to simulate experiments to obtain sufficient data for training the CNN.

The subsurface defect in this work is defined as the internal defect with depth less than one wavelength of laser-generated Rayleigh ultrasonic waves. The central frequencies of the generated Rayleigh ultrasonic waves are about 2.2 MHz, with the corresponding wavelength of 1.36 mm. As the depths of all the subsurface defects are consistent at 1 mm beneath the surfaces, therefore, such defects can be effectively detected by the laser-generated Rayleigh ultrasonic waves. The effectiveness of Rayleigh ultrasonic waves on detection of subsurface defects has been well verified [37], [38].

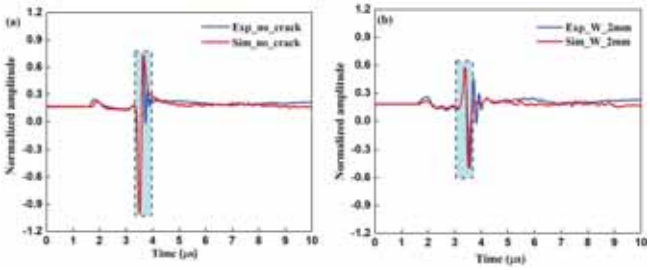

Fig. 4. Comparison between experimental and simulation signals. (a) Signals of no crack (mm). (b) Signals of crack width of 2 mm.

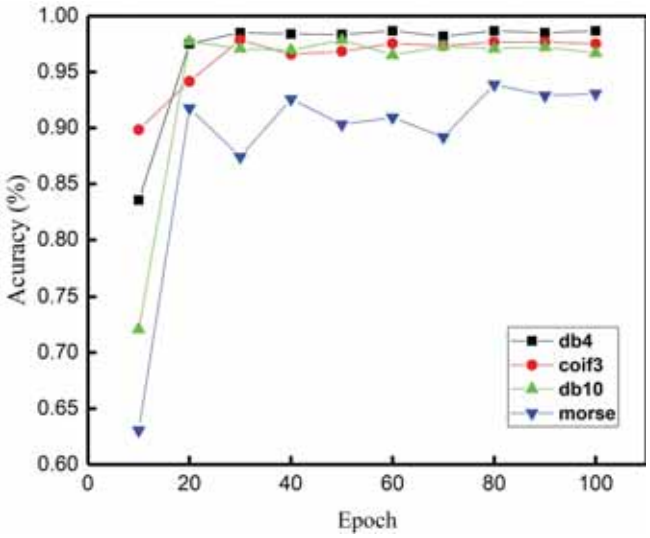

Fig. 5. Validation performance of the models based on various wavelet bases.

B. Wavelet Base Selection and Analysis

For extracting the time–frequency information hidden in the ultrasonic signal more effectively and accurately, wavelet transform is applied to obtain the scalograms. To better represent the features of defect signals, five wavelet bases that are most similar to the waveform of the defect signals are selected from 20 commonly used wavelet bases for subsequent analysis. Fig. 5 shows the comparison of the validation accuracy of the selected wavelet bases. It can be seen that the classification accuracy is affected by the wavelet base and the db4 wavelet base is considered as the most suitable wavelet base to obtain the scalograms.

The interaction of laser-generated ultrasonic waves with various subsurface defects (a crack width of 0.8, 1.6, 2.4, and 3.2 mm) is simulated, and the representative time traces of the laser-generated ultrasonic waves are shown in Fig. 6(a). The result demonstrates that the detected ultrasonic signals are significantly affected by defect width, indicating that the ultrasonic waves encompass the information of crack width. The time-domain signals are transformed into the corresponding scalograms by using the wavelet transform with db4 wavelet base. The variation of crack width significantly affects the corresponding scalogram, in which the pixels characterizing the crack width have conspicuous alteration in color, brightness, and position. Consequently, as input features into the subsequent training model, these db4 scalograms are consid-

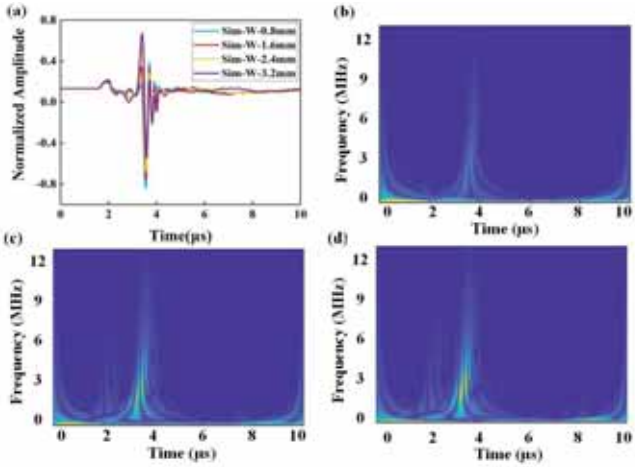

Fig. 6. (a) Simulated time-domain laser ultrasonic signals with various defect widths. (b)–(d) db4 scalograms of ultrasonic signals from specimens with a crack width of 0.8, 1.6, 2.4, and 3.2 mm, respectively.

ered as the optimum images for representing the ultrasonic signal. To enhance the robustness of the later training model, the dataset is enlarged via adding varying levels of white Gaussian noises to the signal of each crack width, which is widely used for training neural network [39].

C. Data Preprocessing

As mentioned above, 31 simulated signals are added with varying levels of white Gaussian noises (SNR ranging from 3 to 30 dB with 3-dB increment), and ten simulations are conducted on each condition. Therefore, the total number of scalogram images in the dataset is 3104 (3100 simulation data and four experimental data), with the randomly selected 80% (2480) simulated images as the training set and remaining 20% (620) simulation images and four experimental data as the validation set. Since the input image for ResNet50 and VGG19 should be of size 224×224, all the images are resized to the target size of 224×224 before training. The batch size for training is set to be 16 so that each epoch in training has 155 steps. Besides, the selections of the optimizer, learning rate, and loss function play a significant role in improving the training effect of the model. In view of the excellent performance of Adam optimizer in many training tasks and its significant advantages such as high computing efficiency and less memory requirement [40], it is selected as the optimizer in this study. The initial learning rate is set to be 0.0003, which is small enough to prevent the parameters directly skipping the local minimum value, resulting in no convergence. The cross-entropy loss is taken as the loss function, as it is more reasonable to update the model parameters in classification tasks.

D. Training Results

The scalograms obtained from the wavelet transform with wavelet base of db4 are used for the model training. As the input of model is images, a great quantity of epochs are required for training to achieve a satisfactory result. Therefore, the model is trained for 100 iterations, and the loss

TABLE II
FOUR METRICS OF THE BEST RESULT ON THE VALIDATION SET

Parameters	Accuracy	Precision	Recall	F1-Score
Value	0.98516	0.98265	0.98415	0.98327

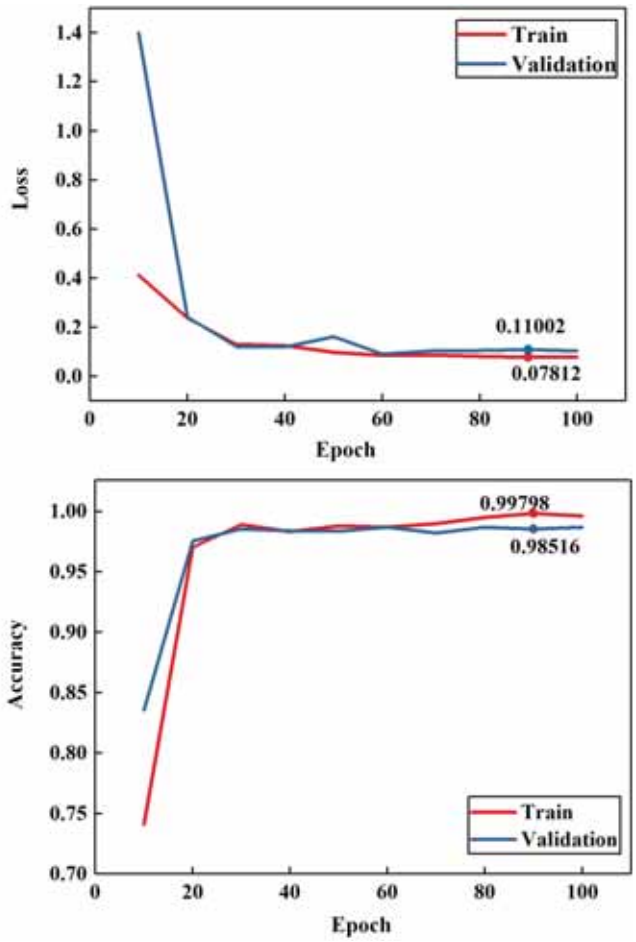

Fig. 7. Accuracy and loss of the model based on db4 scalograms.

value and accuracy are recorded at the end of each iteration. Fig. 7 shows the variations of loss value and accuracy during training and validating, respectively. It can be seen that the validation loss drops dramatically in the first few iterations and then gradually stabilizes at 0.1102, while the training accuracy declines steadily and finally stabilizes at 0.07812, which reveals the effectiveness of the proposed model for training the scalograms. In addition, the tendencies of training accuracy and validation accuracy are very similar, which both converge after 20 iterations of training. Finally, the training accuracy stabilizes at 99.798%, while the overall validation accuracy stabilizes at 98.516%. As the gap between the two is small, which indicates that there is no overfitting problem. Moreover, this approach achieves the prediction accuracy of 100% for the four experimental data, which proves the feasibility and reliability of the established model.

The confused matrix of the best epoch on the validation set is shown in Fig. 8, displaying the amount of correct predictions and error predictions. It shows that CNNs have

Classifier : mLSTM																															
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30	S31	Accuracy
S1 (W=0.0mm)	16	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93.75%
S2 (W=0.2mm)	0	24	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95.83%
S3 (W=0.4mm)	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S4 (W=0.6mm)	0	0	0	20	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95%
S5 (W=0.8mm)	0	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S6 (W=1.0mm)	0	0	0	0	0	14	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92.85%
S7 (W=1.2mm)	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S8 (W=1.4mm)	0	0	0	0	0	0	0	19	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	94.73%
S9 (W=1.6mm)	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S10 (W=1.8mm)	0	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S11 (W=2.0mm)	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S12 (W=2.2mm)	0	0	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S13 (W=2.4mm)	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S14 (W=2.6mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S15 (W=2.8mm)	0	0	0	0	0	0	1	0	0	0	0	0	0	0	26	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92.31%
S16 (W=3.0mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S17 (W=3.2mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S18 (W=3.4mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	100%
S19 (W=3.6mm)	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	0	95.45%
S20 (W=3.8mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	0	0	0	0	1	0	0	0	0	0	95.65%
S21 (W=4.0mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	100%
S22 (W=4.2mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	100%
S23 (W=4.4mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	100%
S24 (W=4.6mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	100%
S25 (W=4.8mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	100%
S26 (W=5.0mm)	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	24	0	0	0	0	0	91.67%
S27 (W=5.2mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	0	0	0	0	100%
S28 (W=5.4mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	100%
S39 (W=5.6mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	0	0	100%
S30 (W=5.8mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	22	0	95.45%
S31(W=6.0mm)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	20	95%
Overall Accuracy : 98.26%																																

Fig. 8. Confusion matrix of the best performance on the validation set.

excellence in defect width prediction, with only minor errors on a few cases. These errors may be caused by the noise and the similar waveform of the defect signals. However, in general, the number of misjudgments in the model is negligible. There are some metrics obtained from the best result on the validation set in Table II, including accuracy, precision, recall rate, and F1-score [41]. All of them indicate that the proposed model has excellent performance on the validation set.

IV. Conclusion

Machine learning has been increasingly utilized in ultrasonic wave-based NDE applications. Conventional approaches necessarily require to select and extract either time- or frequency-domain features from detected ultrasonic signals manually for training the machine learning model. However, this procedure is very subjective, empirical, and time-consuming, and in most cases, it is very difficult to extract the reliable features that are most relevant with the defects to be detected. This work proposes a novel approach combining CNN and wavelet transform to analyze the laser ultrasonic signals for detecting the width of subsurface defects automatically and accurately. The novelty of this work is to convert the laser ultrasonic signals into the scalograms (images) via wavelet transform, which are subsequently utilized as the image input for the pretrained CNN to extract the defect features automatically to quantify the width of defects, avoiding the necessity and inaccuracy induced by artificial feature selection. The experimentally validated numerical model is established to simulate the interaction of laser ultrasonic waves with subsurface defects. The simulated laser ultrasonic signals are converted into scalograms via wavelet transform with the most suitable wavelet base db4. The obtained scalograms representing the time–frequency information of the laser ultrasonic signals are used as the input for training the deep learning model. Due to the advantages of transfer learning, this approach achieves excellent performance with the overall prediction accuracy of 98.5% on the validation set, with the prediction accuracy of 100% for the four experimental data. This work proves that the proposed CNN based on the feature selection approach is reliable and effective to quantify the width of subsurface defects.

The main objective of this work is to prove the feasibility and effectiveness of the proposed approach, and therefore, only the width of defects is investigated. In real applications, as long as sufficient data with regard to the various defect types (dimensions, orientations, locations, and shapes) can be acquired, the developed approach can achieve the accurate detection of corresponding defects.

References

[1] K. Li, Z. Ma, P. Fu, and S. Krishnaswamy, "Quantitative evaluation of surface crack depth with a scanning laser source based on particle swarm optimization-neural network," *NDT &E Int.*, vol. 98, pp. 208–214, Sep. 2018.

[2] K. Zhang, G. Lv, S. Guo, D. Chen, Y. Liu, and W. Feng, "Evaluation of subsurface defects in metallic structures using laser ultrasonic technique and genetic algorithm-back propagation neural network," *NDT &E Int.*, vol. 116, Dec. 2020, Art. no. 102339.

[3] S. Guo, S. Chen, L. Zhang, W. H. Liew, and K. Yao, "Direct-write piezoelectric ultrasonic transducers for pipe structural health monitoring," *NDT &E Int.*, vol. 107, Oct. 2019, Art. no. 102131.

[4] S. Guo, L. Zhang, S. Chen, C. K. I. Tan, and K. Yao, "Ultrasonic transducers from thermal sprayed lead-free piezoelectric ceramic coatings for *in-situ* structural monitoring for pipelines," *Smart Mater. Struct.*, vol. 28, no. 7, Jul. 2019, Art. no. 075031.

[5] S. Guo *et al.*, "Method and analysis for determining yielding of titanium alloy with nonlinear Rayleigh surface waves," *Mater. Sci. Eng., A*, vol. 669, pp. 41–47, Jul. 2016.

[6] D. Lévesque, C. Bescond, M. Lord, X. Cao, P. Wanjara, and J.-P. Monchalin, "Inspection of additive manufactured parts using laser ultrasonics," *AIP Conf. Proc.*, vol. 1706, no. 1. Feb. 2016, Art. no. 130003.

[7] P. Bate, P. Lundin, E. Lindh-Ulmgren, and B. Hutchinson, "Application of laser-ultrasonics to texture measurements in metal processing," *Acta Mater.*, vol. 123, pp. 329–336, Jan. 2017.

[8] H. Selim, M. Delgado, J. Trull, R. Picó, and C. Cojocaru, "Material defect reconstruction by non-destructive testing with laser induced ultrasonics," *J. Phys., Conf.*, vol. 1149, Dec. 2018, Art. no. 012011.

[9] Y. Zeng, X. Wang, X. Qin, L. Hua, and M. Xu, "Laser ultrasonic inspection of a wire+ arc additive manufactured (WAAM) sample with artificial defects," *Ultrasonics*, vol. 110, Feb. 2020, Art. no. 106273.

[10] S. Khan, J. Huh, and J. C. Ye, "Variational formulation of unsupervised deep learning for ultrasound image artifact removal," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 68, no. 6, pp. 2086–2100, Jun. 2021.

[11] R. J. Pyle, R. L. T. Bevan, R. R. Hughes, R. K. Rachev, A. A. S. Ali, and P. D. Wilcox, "Deep learning for ultrasonic crack characterization in NDE," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 68, no. 5, pp. 1854–1865, May 2021.

[12] G. Krummenacher, C. S. Ong, S. Koller, S. Kobayashi, and J. M. Buhmann, "Wheel defect detection with machine learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 1176–1187, Apr. 2018.

[13] G. Yao, T. Lei, and J. Zhong, "A review of convolutional-neural-network-based action recognition," *Pattern Recognit. Lett.*, vol. 118, pp. 14–22, Feb. 2019.

[14] K. Makantasis, A. Doulamis, N. Doulamis, and K. Psychas, "Deep learning based human behavior recognition in industrial workflows," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 1609–1613.

[15] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, "Deep retinal image understanding," in *Proc. Int. Conf. Med. Image Comput. Assist. Intervent.* Cham, Switzerland: Springer, 2016, pp. 140–148.

[16] G. Shi, X. Chen, X. Song, F. Qi, and A. Ding, "Signal matching wavelet for ultrasonic flaw detection in high background noise," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 58, no. 4, pp. 776–787, Apr. 2011.

[17] A. Abbate, J. Koay, J. Frankel, S. C. Schroeder, and P. Das, "Signal detection and noise suppression using a wavelet transform signal processor: Application to ultrasonic flaw detection," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 44, no. 1, pp. 14–26, Jan. 1997.

[18] Z. Ren, K. Qian, Z. Zhang, V. Pandit, A. Baird, and B. Schuller, "Deep scalogram representations for acoustic scene classification," *IEEE/CAA J. Automatica Sinica*, vol. 5, no. 3, pp. 662–669, May 2018.

[19] A. Bhattacharyya, M. Sharma, R. B. Pachori, P. Sircar, and U. R. Acharya, "A novel approach for automated detection of focal EEG signals using empirical wavelet transform," *Neural Comput. Appl.*, vol. 29, no. 8, pp. 47–57, Apr. 2018.

[20] K. Zhang, S. Li, and Z. Zhou, "Detection of disbonds in multi-layer bonded structures using the laser ultrasonic pulse-echo mode," *Ultrasonics*, vol. 94, pp. 411–418, Apr. 2019.

[21] I. Daubechies, *The Wavelet Transform, Time-Frequency Localization and Signal Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 2009.

[22] R. J. Medina-Daza, "Wavelet daubechies (db4) transform assessment for WorldView-2 images fusion," *J. Comput.*, vol. 12, no. 4, pp. 301–308, 2017.

[23] M. Elgendi, M. Jonkman, and F. De Boer, "R wave detection using coiflets wavelets," in *Proc. IEEE 35th Annu. Northeast Bioeng. Conf.*, Apr. 2009, pp. 1–2.

[24] F. B. Vialatte, J. Solé-Casals, J. Dauwels, M. Maurice, and A. Cichocki, "Bump time-frequency toolbox: A toolbox for time-frequency oscillatory bursts extraction in electrophysiological signals," *BMC Neurosci.*, vol. 10, no. 1, pp. 1–12, Dec. 2009.

[25] S. C. Olhede and A. T. Walden, "Generalized morse wavelets," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2661–2670, Nov. 2002.

[26] V. Bostanov, "BCI competition 2003-data sets Ib and IIb: Feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1057–1061, May 2004.

[27] Y. Zhu *et al.*, "Heterogeneous transfer learning for image classification," in *Proc. AAAI Conf. Artif. Intell.*, vol. 11, Aug. 2011, pp. 1304–1309.

[28] R. A. Dunne and N. A. Campbell, "On the pairing of the softmax activation and cross-entropy penalty functions and the derivation of the softmax activation function," in *Proc. 8th Aust. Conf. Neural Netw.*, Melbourne, FL, USA, vol. 181, Jun. 1997, p. 185.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

[31] S. Rajaraman *et al.*, "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images," *PeerJ*, vol. 6, p. e4568, Apr. 2018.

[32] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.

[33] Y. Hold-Geoffroy, A. Athawale, and J.-F. Lalonde, "Deep sky modeling for single image outdoor lighting estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 6927–6935.

[34] K. S. Shergill, K. S. Banu, and B. Tripathy, "An improved differential neural computer model using multiplicative LSTM," in *Proc. Int. Conf. Soft Comput. Syst.* Singapore: Springer, 2018, pp. 283–290.

[35] B. Krause, L. Lu, I. Murray, and S. Renals, "Multiplicative LSTM for sequence modelling," 2016, *arXiv:1609.07959*. [Online]. Available: https://arxiv.org/abs/1609.07959

[36] G. Du, L. Bu, Q. Hou, J. Zhou, and B. Lu, "Prediction of the compressive strength of high-performance self-compacting concrete by an ultrasonic-rebound method based on a GA-BP neural network," *PLoS ONE*, vol. 16, no. 5, May 2021, Art. no. e0250795.

[37] C. Wang, A. Sun, X. Yang, B.-F. Ju, and Y. Pan, "Numerical simulation of the interaction of laser-generated Rayleigh waves with subsurface cracks," *Appl. Phys. A, Solids Surf.*, vol. 124, no. 9, pp. 1–10, Sep. 2018.

[38] A. Kromine, P. Fomitchov, S. Krishnaswamy, and J. Achenbach, "Detection of subsurface defects using laser based technique," *AIP Conf. Proc.*, vol. 557, no. 1, pp. 1612–1617, 2001.

[39] A. Neelakantan *et al.*, "Adding gradient noise improves learning for very deep networks," 2015, *arXiv:1511.06807*. [Online]. Available: https://arxiv.org/abs/1511.06807

[40] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980

[41] A. Korotcov, V. Tkachenko, D. P. Russo, and S. Ekins, "Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets," *Mol. Pharmaceutics*, vol. 14, no. 12, pp. 4462–4475, Dec. 2017.

Shifeng Guo received the M.S. degree in mechanical engineering from Chongqing University, Chongqing, China, in 2008, and the Ph.D. degree in mechanical engineering from the National University of Singapore (NUS), Singapore, in 2013.

He is currently a Professor with the Shenzhen Institutes of Advanced Technology (SIAT), Chinese Academy of Sciences, Shenzhen, China. He has authored or coauthored over 60 peer-reviewed journal publications and over 20 invited talks in his research fields. His research mainly focuses on flexible sensors and sensor networks, high-resolution ultrasonic imaging, ultrasonic nondestructive evaluation and structural health monitoring, smart materials and structures, micromechanics–nanomechanics, and machine learning.

Haowen Feng received the B.Sc. degree from the Guangdong University of Foreign Studies, Guangzhou, China, in 2019. He is currently pursuing the master's degree with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China.

His research interests include laser ultrasonic and machine learning.

Wei Feng received the B.Sc. and Ph.D. degrees from the Huazhong University of Science and Technology, Wuhan, China, in 2001 and 2006, respectively.

He is currently a Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His main research interests include digital intelligent manufacturing and the Industrial Internet.

Gaolong Lv received the M.S. degree in mechanical engineering from Shandong Agricultural University, Tai'an, China, in 2018.

He is currently a Research Assistant with the Shenzhen Institutes of Advanced Technology (SIAT), Chinese Academy of Sciences, Shenzhen, China. His research interests include laser ultrasonic, signal processing, and machine learning.

Dan Chen received the B.E. and Ph.D. degrees from the University of Science and Technology Beijing, Beijing, China, in 2013 and 2019, respectively.

He is currently an Assistant Researcher with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research interests include laser ultrasonics, signal processing, and photoacoustic transducer.

Yanjun Liu received the B.E. degree in optoelectronics from Shandong University, Jinan, China, in 2000, the M.S. degree in optics from Fudan University, Shanghai, China, in 2003, and the Ph.D. degree in photonics from Nanyang Technological University, Singapore, in 2007.

He is currently an Associate Professor with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology (SUSTech), Shenzhen, China. He has authored or coauthored over 170 peer-reviewed journal publications and over 20 invited talks in his research fields. His research group mainly focuses on liquid crystal photonics, active plasmonics, and metamaterials.

Xinyu Wu (Member, IEEE) received the B.E. and M.E. degrees from the Department of Automation, University of Science and Technology of China, Hefei, China, in 2001 and 2004, respectively, and the Ph.D. degree from The Chinese University of Hong Kong, Hong Kong, in 2008.

He is currently a Professor with the Shenzhen Institutes of Advanced Technology and the Director of the Center for Intelligent Bionic, Chinese Academy of Sciences, Shenzhen, China. He has published over 180 articles and two monographs. His research interests include computer vision, robotics, and intelligent systems.