Ultrasonic Assessment of Thickness and Bonding Quality of Coating Layer Based on Short-Time Fourier Transform and Convolutional Neural Networks

Azamatjon Kakhramon ugli Malikov 1, Younho Cho 1,*, Young H. Kim 2, Jeongnam Kim 1, Junpil Park 2 and Jin-Hak Yi 3

1 School of Mechanical Engineering, Pusan National University, Busan 46241, Korea; malikov@pusan.ac.kr (A.K.u.M.); kjnnoah@pusan.ac.kr (J.K.)
2 Institute of Nuclear Safety and Management, Pusan National University, Busan 46241, Korea; yhkim627@pusan.ac.kr (Y.H.K.); jpp@pusan.ac.kr (J.P.)
3 Coastal Development and Ocean Energy Research Center, Korean Institute of Ocean Science & Technology, Busan 49111, Korea; yijh@kiost.ac.kr
* Correspondence: mechcyh@pusan.ac.kr; Tel.: +82-51-510-2323

Abstract: Ultrasonic non-destructive analysis is a promising and effective method for the inspection of protective coating materials. Offshore coating exhibits a high attenuation rate of ultrasonic energy due to the absorption and ultrasonic pulse echo testing becomes difficult due to the small amplitude of the second echo from the back wall of the coating layer. In order to address these problems, an advanced ultrasonic signal analysis has been proposed. An ultrasonic delay line was applied due to the high attenuation of the coating layer. A short-time Fourier transform (STFT) of the waveform was implemented to measure the thickness and state of bonding of coating materials. The thickness of the coating material was estimated by the projection of the STFT into the time-domain. The bonding and debonding of the coating layers were distinguished using the ratio of the STFT magnitude peaks of the two subsequent wave echoes. In addition, the advantage of the STFT-based approach is that it can accurately and quickly estimate the time of flight (TOF) of a signal even at low signal-to-noise ratios. Finally, a convolutional neural network (CNN) was applied to automatically determine the bonding state of the coatings. The time–frequency representation of the waveform was used as the input to the CNN. The experimental results demonstrated that the proposed method automatically determines the bonding state of the coatings with high accuracy. The present approach is more efficient compared to the method of estimating bonding state using attenuation.

Keywords: coating with high attenuation; thickness and bonding status; ultrasonic pulse-echo; short-time Fourier transform (STFT); convolutional neural networks (CNN)

1. Introduction

Coatings are widely applied to offshore structures as corrosion protection layers. Coating layer thinning or lack of adhesion between the coating layer and structure may affect the safety of the structure and ignoring this can result in complete system failure. The lifetime of the coated structure is directly related to the quality of the adhesion of the coating material with the structure. The thickness and adhesion state are two important features of coatings that must be monitored periodically. The thickness of the coating layer relies on the surrounding environment of the structure and the thickness of coatings may vary in different fields of application. In practical applications, the coating is significantly thin and the thickness may vary up to 3 mm. However, in some applications, the coating material is applied as a thin layer on the surface of the metallic structure. In such cases, the thickness of the coating material may be close to 1 mm or even thinner. In the field, the coating material is usually applied by conventional rollers [1] that create many pores...
on the inside of the polymer. Porosity greatly increases the ultrasonic attenuation of the coating layer, thus ultrasonic testing becomes difficult.

Ultrasound are widely used in the engineering field and can be divided into the cases of using energy and wave propagation. High intensity ultrasound using energy is used in many fields such as cleaning, material processing [2–4], and medical therapy [5]. In this case, controls such as initial conditions is key. Ultrasonic measurement using wave propagation is also used in many fields such as non-destructive inspection of structures, quality control of products, sonar, and medical diagnosis. In this case, precise measurement of propagation time and amplitude is key.

Ultrasonic non-destructive technology is widely used in flaw detection and materials evaluation. Piezoelectric transducers that transmit and receive ultrasonic waves have a dead zone called the “main bang” that establishes the pulse–echo test method as unable to detect fast returning echoes [6]. In the case of short propagation times, the reflected echoes are superimposed or buried in the dead zone [7–9]. Transducers with higher center frequencies have shorter pulse durations and can be used to solve echo-overlapping problems [10]. Conversely, as high-frequency energy is rapidly attenuated, it cannot be applied to the evaluation of the thickness of the coating layer with large attenuation [11].

The time of flight (TOF) of the echo reflected from the layer boundary is required for ultrasonic pulse–echo testing. The back wall echoes of coating arrive very close to each other and accurate estimation of the TOF of an echo signal is a key issue. Studies have been conducted on the Morlet wavelet and least-squares filters [12], the combination of the Hilbert transform and the non-linear least-squares method [13], the principle of maximum likelihood estimation [14] and the cross-correlation method [15,16], and the ellipse method [17].

The ultrasonic non-destructive evaluation of the state of bonding is particularly important to ensure the long-term protection of coated structures. For this purpose, studies regarding, for example, echo signal parameters and joint strength [18], asphalt separation detection using surface waves [19], the ultrasonic C-scan method [20], and the correlation between guided wave amplitudes [21] have been conducted. In addition, the phase of the echo signal and synthetic aperture focusing technique imaging algorithm [22–24], the longitudinal wave velocity variation [25,26], the reflected wave velocity frequency characteristic [27,28], the combination of time-domain waveform and numerical simulation [29,30], the support matching method [31], and a method for comparing damping rates [32] are also areas that have been studied.

The STFT technique splits the signal into multiple sections using a sliding window function along the time axis and Fourier transform is applied to each section of the signal [33]. Several researchers have presented results in the evaluation of the bonding state of the adhesive layer using the STFT. For example, Song et al. [34] evaluated the bonding level between shotcrete and planar rock using STFT analysis with an impact echo test [35]. In contrast, in this study, STFT was used both for the precise thickness estimation and bonding state evaluation using the ultrasonic pulse–echo method.

Finally, deep learning algorithms such as convolutional neural networks (CNNs) have been implemented to enable the fast automated identification of the bonding state. Manual inspection of the state of the coating layer is time consuming, expensive, and prone to errors. The detection of the coating layer relies on the amplitude ratio of the ultrasonic wave energy and the presence of the reflected echo signal from the base material. In spite of the significant achievements in the non-destructive inspection of coating layers, most of research results are qualitative and not quantitative. The attenuation-based evaluation of the coating layer presents challenges due to the presence of non-uniform porosity in the layer that causes a variation in the attenuation parameters [36,37].

The purpose of the present study is to develop an advanced ultrasonic assessment system for offshore coatings. An ultrasonic delay line was adopted to separate the reflected back wall echo from the excitation signal (main bang). A short-time Fourier transform (STFT)-based analysis of the waveform has been implemented to estimate the coating layer
Coatings 2021, 11, 909

thickness and bonding state of the coating materials. In addition, STFT spectrogram is integrated with a CNN to provide an automatic classification of the bonding condition of the coating layer. The accuracy and performance of the developed system were evaluated.

2. Short-Time Fourier Transform (STFT)

STFT is a powerful method for analyzing non-stationary signals. Conventional Fourier transform of the signal provides averaged frequency of the signal over an entire time interval. In contrast to the standard Fourier transform, the STFT allows for the representation of the frequency variation of the signal over time. When the STFT method is applied, the signal is divided into equal length sections by the windowing function and the Fourier transform is applied to each section. Using STFT, the continuous time signal \( x(u) \) is windowed with a windowing function \( h(u) \) to a limited extent and then Fourier transforms are applied. The expression of STFT [38] is illustrated by

\[
STFT(t, f) = \int_{-\infty}^{+\infty} x(u) h^*(u - t) e^{-j2\pi fu} \, du
\]

The windowing function is a non-negative and even function with limited duration. In Equation (1), \( h^*(u) \) is the complex conjugate of the windowing function. The windowing function attempts to reach the peak of the signal and emphasizes it at that instant, while the remaining section of the signal is suppressed by the same window function. Several kinds of window functions exist in the literature: consider Kaiser, Blackman, Gaussian, Nuttall, Chebyshev, and others [39,40]. In the current study, the Gaussian windowing function was applied to STFT analysis and can be defined as [41]

\[
h(t) = \left( \frac{1}{\sqrt{2\pi a}} \right) e^{-\frac{1}{2}(\frac{t}{a})^2}
\]

Gaussian windowing has advantages for the time–frequency representation of the waveform due to its optimal concentration in both the time and frequency domains. The scale factor (or window duration) \( a \) in Equation (2) determines the size or duration of the windowing function. The scale factor of the windowing function must be selected carefully because it determines the resolution of both the time and frequency domains. An infinitely fine resolution of time and frequency cannot be obtained simultaneously. To achieve good frequency resolution, a long windowing function is required, whereas a better time resolution is obtained with a short windowing function. The duration of the Gaussian windowing function for optimal resolution of both in time and frequency domain can be computed by [42]

\[
a = \sqrt{\frac{\lambda}{2\pi^2}}
\]

The reflected echo from the back wall can be described by a Gaussian echo model [42]. The back wall echo model is described in Equation (4) and the parameters of the equation are the following: \( \beta \) is the amplitude, \( \lambda \) is the bandwidth factor, \( f_c \) is the center frequency, \( \tau \) is the TOF, and \( \phi \) is the phase of the reflected echo signal.

\[
x(t) = \beta e^{-(\lambda(t-\tau)^2+j(2\pi f_c(t-\tau)+\phi))}
\]

The magnitude of the STFT can be calculated by substituting the model of the back wall echo (Equation (4)) into the STFT expression (Equation (1)) in addition to the Gaussian windowing function (Equation (2)). The simplified expression of the magnitude of the STFT can be defined as [41]

\[
\left| STFT(t, f) \right| = \frac{\beta}{\sqrt{2\lambda a^2 + 1}} \exp\left[-\frac{\lambda}{2\lambda a^2 + 1} (t - \tau)^2 - \frac{2\pi^2 a^2}{2\lambda a^2 + 1} (f - f_c)^2 \right]
\]
As evident in Equation (5), the expression of the magnitude of the STFT of the back wall echo model is a function of both time and frequency. The STFT can be represented as a three-dimensional plot where the x, y, and z axes correspond to the time, frequency, and intensity of the STFT magnitude, respectively. Although, useful information can be derived from a two-dimensional view of the STFT (i.e., time–frequency or time–magnitude) as well. By using the obtained magnitude of the STFT, the waveform parameters such as the TOF and center frequency can be estimated. To calculate the TOF of the waveform, the obtained STFT magnitudes must be projected into the time-domain and the local peaks of the projected STFT magnitude into the time-domain \( \frac{\partial |\text{STFT}(t,f)|}{\partial t} = 0 \) correspond to the TOF of the waveform. Similarly, the center frequency of the waveform can be evaluated by estimating \( \frac{\partial |\text{STFT}(t,f)|}{\partial f} = 0 \) of the STFT projection to the frequency domain which corresponds to the central frequency [43].

The magnitude of the STFT shows the distribution of the energy of the waveform over the time–scale plane. The waveform energy \( E \) (Equation (6)) distributed over the time–frequency plane can be obtained by integrating the magnitude of the STFT and the relation can be defined as [44–46]

\[
E = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\text{STFT}(t,f)|^2 dt df
\]

3. Experiments

3.1. Ultrasonic Echoes from the Coating Layer

Ultrasonic delay lines were used to shift the first echo in the “main bang”. Figure 1 shows the wave propagation patterns of fully bonded and debonded coatings when delay lines are used, where A0 represents the echo reflected between the delay line and the coating, and A1 represents the echo reflected from the back wall of the coating layer. When the coating layer adheres well to the base material, part of the ultrasonic energy is propagated into the base material; thus, the amplitude of the A1 echo becomes smaller than when the coating layer and the base material are separated. When the coating layer is bonded with carbon steel, the A2 corresponds to the back wall echo waveform from the base material, whereas in the debonded layer, A2 is the second echo from the back wall of the coating material.

![Figure 1. Schematics of the waveform path: (a) perfectly bonded and (b) debonded case between the coating and base material.](image)

In this study, the case of debonding and bonding between the coating material and structure was distinguished using the amplitude ratio of the peak STFT magnitude. In reference to Equation (5), the STFT magnitude peak is proportional to the waveform amplitude \( \beta \) [47]; thus, the ratio of the STFT magnitude peaks of the two subsequent wave echoes was used to compare the amplitude ratio of the two waveforms.

3.2. Test Sample Preparation

To validate the proposed method in actual experiments, specimens were fabricated using two types of polymer coating materials. Both coating materials are composed of
an epoxy resin and a hardener, and detailed information on the two coating materials is presented in Table 1 [48]. Carbon steel plates with dimensions of 150 mm × 50 mm × 4 mm were used as the base material for both types of coatings.

Table 1. Information/Material Properties of the coating materials.

<table>
<thead>
<tr>
<th>Information/Properties</th>
<th>Epoxy A</th>
<th>Epoxy B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial name</td>
<td>ALOCIT 28.14 RS</td>
<td>RS 500P</td>
</tr>
<tr>
<td>Maker</td>
<td>A&amp;E Group Chemco</td>
<td>Chemco</td>
</tr>
<tr>
<td>Weight ratio</td>
<td>5:1</td>
<td>5.1:1</td>
</tr>
<tr>
<td>(Epoxy resin: hardener)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Density</td>
<td>1.94 g/cm³</td>
<td>1.67 g/cm³</td>
</tr>
<tr>
<td>Ultrasonic wave speed</td>
<td>2.46 mm/µs</td>
<td>2.40 mm/µs</td>
</tr>
</tbody>
</table>

The epoxy resin and hardener for each coating were mechanically mixed for 3–5 min in separate plastic containers. There are several ways to apply the coating [49] but as is common in the field, it was applied individually to the surface of the base material using conventional rollers. As hand mixing and roller application are used, there are many pores inside the coating layer. They increase attenuation in the coating layer and thus ultrasound examination is difficult.

The surfaces of the carbon steel plates were divided into two parts. In the first part, coating material was directly applied to the surface of the base materials to create a perfect bond between the carbon steel and the coating. In the second part, a thin sheet of stainless steel with a thickness of 0.05 mm was placed between the coating and the base material, and after curing the coated specimen at room temperature, the thin stainless steel sheet was removed. As a result, the coating layer and carbon steel were completely separated and a completely debonded layer of the coating was obtained. The fabricated specimens are shown in Figure 2.

Figure 2. Fabricated specimens of the coating layers: (a) epoxy A (b) epoxy B.

3.3. Experimental Set-Up

Figure 3 illustrates the instrune setup used for the ultrasonic pulse–echo measurements of the specimens. The pulser-receiver (5072 PR, Panametrics, Waltham, MA, USA) was used for the generation and reception of the ultrasound. The digital oscilloscope (WaveRunner 604Zi, Teledyne LeCroy, Chestnut Ridge, NY, USA) was used for averaging and recording the received waveforms. A piezoelectric transducer with a central frequency of 5 MHz was used. A lower frequency transducer generated long-lasting signals that led to the overlapping of the echo waves, whereas at high frequencies, the amplitude ratio was increased. Also, an acrylic delay line of 6.2 mm thick was used.
4. Results and Discussions

4.1. Coating Layer Thickness Measurement

All thickness measurements were performed on the debonded area of the coating layer. The main purpose of this experiment was to measure the thickness of the coating layer and as mentioned in Section 2, the TOF of the wave packet can be determined by the local maximum value of projection of the STFT magnitude into the time-domain. Thickness measurement experiments were performed on coating materials of various thicknesses. In the first step, coatings with a thickness of 1.1–1.5 mm were examined using ultrasound. Figures 4 and 5 show the experimental results from the direct pulse–echo test of the epoxy A and epoxy B coating layers, respectively.

The peaks of the STFT magnitude can be observed in the figures in which the first echo denoted by A₀ is the signal reflected from the boundary between the delay line and the coating layer surface, and the echo denoted by A₁ is the signal reflected from the boundary between the base material and the coating material. By knowing the wave velocity in the coating and the TOF difference of the two echoes, the thickness of the coating layer is calculated.

The difference in acoustic impedance of the delay line and the coating material causes the larger magnitude of the A₁ echo compared to the A₀, as shown in Figure 5. More precisely, the acoustic impedances of epoxy A and B are $4.77 \times 10^6$ kg/(m²s) and $4.01 \times 10^6$ kg/(m²s), respectively, which is similar to the acoustic impedance of the acrylic...
delay line \( (3.26 \times 10^6 \text{ kg/(m}^2\text{s)}) \). This is because the transmitted energy is greater than the reflected energy at the interface.

![Figure 5. Epoxy B coating (thickness less than 1.5 mm): (a) time-domain waveform and (b) the projection of the STFT.](image)

In epoxy A, the longitudinal wave velocity is 2.46 mm/µs, the TOF of the first wave echo obtained in Figure 4b is 5.28 µs, and the arrival time of the next echo is 6.40 µs; thus, the thickness of the coating layer is 1.38 mm. Similarly, the longitudinal wave velocity of epoxy B is 2.40 mm/µs and the corresponding TOFs for echoes A0 and A1 are 5.28 µs and 6.31 µs, respectively; thus, the thickness was estimated to be 1.24 mm.

The experiment was repeated with thinner layers of the epoxy A and B coatings and the STFT-based method was applied to the obtained waveforms; the results are shown in Figures 6 and 7. By the above method, the thicknesses of the epoxy A and B coatings were estimated to be 0.91 mm and 0.92 mm, respectively.

![Figure 6. Pulse-echo test of the epoxy A coating layer with a thickness less than 1 mm: (a) time-domain waveform and (b) the projection of the STFT of the waveform.](image)
Figure 7. Epoxy B coating (thickness less than 1 mm): (a) time-domain waveform and (b) the projection of the STFT.

Measurement data using ultrasound was verified by a manual measurement using calipers. Repeatability was checked by measuring five times the same thickness by each method and the standard deviation (std dev.) was calculated; values are shown in Table 2. Comparing the thickness values measured with both methods showed similar results and the difference between the two values was about 1%.

Table 2. Coating layer thickness measurements by caliper and ultrasound.

<table>
<thead>
<tr>
<th>Coating Material</th>
<th>Thickness Measured by Caliper</th>
<th>Thickness Measured by Ultrasound</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Average)         (Std Dev.)</td>
<td>(Average)         (Std Dev.)</td>
<td></td>
</tr>
<tr>
<td>Epoxy A</td>
<td>1.37 mm            0.0074</td>
<td>1.38 mm            0.0040</td>
<td>0.7%</td>
</tr>
<tr>
<td>Epoxy A</td>
<td>0.92 mm            0.0075</td>
<td>0.91 mm            0.0070</td>
<td>1.1%</td>
</tr>
<tr>
<td>Epoxy B</td>
<td>1.25 mm            0.0040</td>
<td>1.24 mm            0.0108</td>
<td>0.8%</td>
</tr>
<tr>
<td>Epoxy B</td>
<td>0.91 mm            0.0109</td>
<td>0.92 mm            0.0045</td>
<td>1.1%</td>
</tr>
</tbody>
</table>

The thickness of the coating layer was measured by applying the STFT which was able to accurately measure the thickness of less than 1 mm for both types of coating materials. Although the TOF of the reflected echo can be evaluated from the time-domain waveform, the advantage of the STFT-based approach is that it can accurately and quickly estimate the TOF of the signal even at low signal-to-noise ratios.

4.2. Evaluation of Bonding Status

In order to evaluate the bonding state between the coating layer and the base material, ultrasonic experiments were performed on the bonded and debonded parts, and the results are shown in Figures 8 and 9, respectively. The received time–domain waveform is presented in (a) and (b) presents the time–frequency representation obtained using STFT which is used as an input signal for the CNN that will be mentioned later. In addition, (c) presents the magnitude of the projected STFT in the time-domain, allowing for the accurate estimation of the TOF of the echo. Similarly, Figures 10 and 11 present the results measured for the perfectly bonded and debonded portions of the epoxy B specimen, respectively.
Figure 8. Ultrasonic signal from the perfectly bonded sections of the epoxy A: (a) time-domain waveform, (b) time-frequency representation of the waveform, and (c) the projection of STFT of the waveform into the time-domain.

Figure 9. Experimental result of the debonded epoxy A coating layer: (a) time-domain signal, (b) spectrogram, and (c) the STFT magnitude projection.
Figure 10. Ultrasonic signal from the pulse-echo experimental result of the perfectly bonded section of epoxy B: (a) time-domain waveform, (b) time-frequency representation, and (c) the projection of the STFT.

Figure 11. Debonded layer of epoxy B coating: (a) time-domain waveform, (b) spectrogram, and (c) the STFT magnitude projection.

Parameters of the received waveform such as the TOF and magnitude of the STFT of bonded sections of epoxy A and B are listed in Tables 3 and 4. The coating layer has relatively high damping due to the presence of porosity inside but the base material is carbon steel with a low damping rate. At the interface between the coating and the substrate, part of the ultrasound is transmitted from the coating to the substrate and the remainder is reflected back towards the ultrasonic transducer (Echo A1). Ultrasonic waves transmitted to the base material return to the transducer after being reflected several times in the base material (Echo A2 and A3).
Table 3. Estimated properties of the signal from a perfectly bonded and debonded layer of epoxy A.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfectly Bonded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A0</td>
<td>$t_1 = 5.26$</td>
<td></td>
<td>0.21</td>
</tr>
<tr>
<td>A1</td>
<td>$t_2 = 6.02$</td>
<td>$\Delta t_{21} = 0.76$</td>
<td>0.48</td>
</tr>
<tr>
<td>A2</td>
<td>$t_3 = 7.35$</td>
<td>$\Delta t_{32} = 1.33$</td>
<td>0.07</td>
</tr>
<tr>
<td>A3</td>
<td>$t_4 = 8.68$</td>
<td>$\Delta t_{43} = 1.33$</td>
<td>0.06</td>
</tr>
<tr>
<td>Debonded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A0</td>
<td>$t_1 = 5.26$</td>
<td></td>
<td>0.21</td>
</tr>
<tr>
<td>A1</td>
<td>$t_2 = 6.02$</td>
<td>$\Delta t_{21} = 0.76$</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Table 4. Estimated properties of the signal from a perfectly bonded and debonded coating layer of epoxy B.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfectly Bonded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A0</td>
<td>$t_1 = 5.26$</td>
<td></td>
<td>0.17</td>
</tr>
<tr>
<td>A1</td>
<td>$t_2 = 6.09$</td>
<td>$\Delta t_{21} = 0.83$</td>
<td>0.47</td>
</tr>
<tr>
<td>A2</td>
<td>$t_3 = 7.42$</td>
<td>$\Delta t_{32} = 1.33$</td>
<td>0.11</td>
</tr>
<tr>
<td>A3</td>
<td>$t_4 = 8.75$</td>
<td>$\Delta t_{43} = 1.33$</td>
<td>0.09</td>
</tr>
<tr>
<td>Debonded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A0</td>
<td>$t_1 = 5.26$</td>
<td></td>
<td>0.17</td>
</tr>
<tr>
<td>A1</td>
<td>$t_2 = 6.09$</td>
<td>$\Delta t_{21} = 0.83$</td>
<td>0.60</td>
</tr>
</tbody>
</table>

In the case of the debonded layer, air exists between the base material and the coating film, and almost all of the ultrasonic waves are reflected from the back wall of the coating and reach the ultrasonic transducer (Echo A1).

The amplitude of the first echo (A0) which is reflected off the back wall of the delay line is the same for both the bonded and debonded conditions. The bond between the coating layer and the base material transmits a certain amount of ultrasonic energy from the coating material to the base material, reducing the size of the A1 echo. Therefore, the ratio of the two echoes (A1/A0) can be used to evaluate the bonding state of the coating layer. In the case of a perfectly bonded specimen, $A1/A0 = 2.28$, which is smaller than $A1/A0 = 2.95$ if not bonded. Similarly, in the case of the epoxy B specimen, the bonded specimen is $A1/A0 = 2.76$, which is smaller than $A1/A0 = 3.53$ in the case of the unbonded specimen.

Goglio et al. obtained similar results [50]. They observed that the rate of amplitude decay of the ultrasonic echoes was faster in layer with adhesions compared to the layer without adhesions. However, ultrasonic detection of the debonding state of the coating layer must consider complex situations such as delay lines, acoustic impedance mismatch between the coating material and the base materials, high attenuation of the coating material, various thicknesses of the coating layer, and waveform mixing between reflected and transmitted echoes. In addition, the ultrasonic echo ratio comparison technique is performed manually and the peak ratio of the STFT magnitudes depends on several parameters of the material properties of both the coating and the delay line. To overcome this, a CNN was used to automatically and quickly detect the bonding state of the coating layer.

5. Artificial Neural Network

We propose a combined technique of the STFT spectrogram with a CNN for automatically classifying the state of the coating layers. CNNs are a class of deep learning networks [51]. In contrast to traditional neural networks, CNNs contain feature extracting convolutional and pooling layers with a local connection. Direct implementation of neural
networks for coating material classification can lead to the usage of many parameters and is therefore a time-consuming training process [52].

5.1. Structure of CNN

The structure of the CNN consists of several layers including the convolutional layer, activation layer, pooling layer, and fully connected layer [53]. Figure 12 shows the architecture of the CNN used in this study. As can be observed from the figure, the CNN structure consists of convolution, pooling, and fully connected layers. Initially, the time–frequency representation of the waveform is computed using STFT and then the 2D image of the spectrogram is used as an input to the convolutional layer of the CNN. The hyperparameters that determine the CNN structure are classified into two types: network structure hyperparameters and trained network hyperparameters. The kernel type, kernel size, padding, stride, and activation function correspond to network structure parameters. The training network parameters are learning rate, momentum, epoch number, and batch size.

The convolution layer consists of kernel filters that require significant computational resources compared with other layers. In our model, the convolutional layer had a total of 64 filters with dimensions of $3 \times 3$. Low resolution images are formed by kernel filters that are called a feature map. As an activation function of the convolution layer, a rectified linear unit (ReLU) was used. The applied ReLU activation function is to normalize the input image and increase the non-linearity of the convolutional layer output. The output of the ReLU function is zero when it receives a zero or negative value, while the positive values are kept unchanged. The ReLU activation function is defined as [54]

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$$

In the pooling layer, max-pooling was applied with a kernel size of $2 \times 2$ and the stride was equal to two. The feature maps are further downsized by the pooling layer and then were flattened. In the flattening process, all matrices are reordered in a column and were used as the input of the fully connected layer. The fully connected layer consists of several neurons and each neuron generates a single output (0 or 1) while having single or multiple inputs. The input of the neuron is associated with a weight and its value is directly related to the importance of the input value. In the designed CNN, the output of the fully connected layer belonged to a single neuron that was based on the sigmoid activation function. Because the applied CNN model output required to be binary classified the state of bonding of the coating layer whether to bonded or debonded condition of the layer.

The original spectrogram image has $768 \times 1536$ pixels. Spectrogram images were downsized to $128 \times 256$ pixels and were applied for the training of the CNN. Table 5

![Figure 12. CNN architecture consisting of single convolution, pooling, and fully connected layers.](image-url)
summarizes the details of the output feature sizes of the CNN layers. During the training of the CNN, the Adam (adaptive moment estimation) algorithm was implemented as an optimizer and the learning rate was set to 0.003. The binary cross-entropy function was used as a loss function. The parameter of the CNN is summarized in Table 6. The parameters of the CNN could be further optimized until optimal values are reached by using boosting algorithms. In the literature, three several known boosting methods are presented [55]: AdaBoost [56], XGBoosting [57], and Gradient boosting [58], among others. For example, Bustillo et al. [59] implemented AdaBoost ensembles to optimize the parameters of neural networks for data extraction features from tests and to improve the accuracy of prediction to 5%. In general, when the prediction accuracy by a neural network is low, the parameters of the neural network are optimized, and the optimization can affect the overall accuracy and performance of the neural networks. However, optimization parameters of the CNN require significant computation resources and optimal parameter settings that are beyond the scope of this study.

Table 5. Specification of the model architecture.

<table>
<thead>
<tr>
<th>Layer Type</th>
<th>Output Shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input layer</td>
<td>128 × 256 × 3</td>
</tr>
<tr>
<td>Convolution layer</td>
<td>126 × 254 × 64</td>
</tr>
<tr>
<td>Max pooling</td>
<td>63 × 127 × 64</td>
</tr>
<tr>
<td>Flattening</td>
<td>512,064</td>
</tr>
<tr>
<td>Fully connected layer</td>
<td>512</td>
</tr>
<tr>
<td>Predictions</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6. CNN structure hyperparameters.

<table>
<thead>
<tr>
<th>Hyperparameters of the CNN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning algorithm</td>
<td>Adam</td>
</tr>
<tr>
<td>Non-linearity activation function</td>
<td>ReLU</td>
</tr>
<tr>
<td>Activation function</td>
<td>Sigmoid</td>
</tr>
<tr>
<td>Number of epochs</td>
<td>30</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.003</td>
</tr>
<tr>
<td>Number of filters</td>
<td>64</td>
</tr>
<tr>
<td>Kernel size</td>
<td>3 × 3</td>
</tr>
<tr>
<td>Batch size</td>
<td>32</td>
</tr>
<tr>
<td>Momentum</td>
<td>0.9</td>
</tr>
</tbody>
</table>

The main advantage of the CNN concerns the presence of a feature extracting layer in the structure. This allows for the direct use of the spectrograms as the input of the CNN without any additional processing. However, the disadvantage of the CNN concerns the large number of hyperparameters that significantly affect the performance of CNNs. Apart from the many algorithms of optimizing hyperparameters, there is a limit to establishing an optimal combination of all parameters of the CNN. In most cases, main hyperparameters are set manually by trial and error.

5.2. Results of the CNN

In a method to classify the bonding state of a coating material based on the amplitude of the ultrasonic echo, the peak ratio of the STFT magnitudes depends on several parameters of the material properties of the coating and delay lines. To solve this problem, the CNN was developed to automatically and quickly classify the bonding state of the coating layer. The Keras library was used to build the proposed CNN model and it was applied to both types of coating materials. In addition, ultrasonic measurements were recorded on several samples of bonded and debonded sections of both types of coating materials. A total of 4400 ultrasonic inspections were performed on both coating materials (epoxies A and B) with a uniform thickness of the coating layer. Among overall measurements, 2000 were conducted
on the bonded sections of the specimens and the remaining 2200 tests were conducted on the completely debonded sections. The spectrogram of each measured waveform was calculated by applying STFT and 2D image of spectrogram was used as an input of the CNN model. The spectrogram representation was implemented in a subset for the training, validation, and testing of the CNN model. The CNN model was trained separately for the epoxy A and epoxy B type of coating for 30 epochs and individual performance was evaluated for each type of coating material. The sizes of the used subset of both the epoxy A and epoxy B types of materials are listed in Table 7.

Table 7. Distribution of the dataset.

<table>
<thead>
<tr>
<th>Bonding condition</th>
<th>Training set</th>
<th>Validation set</th>
<th>Test set</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfectly bonded</td>
<td>1000</td>
<td>400</td>
<td>600</td>
<td>2000</td>
</tr>
<tr>
<td>Debonded</td>
<td>1200</td>
<td>400</td>
<td>600</td>
<td>2200</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bonding condition</th>
<th>Training set</th>
<th>Validation set</th>
<th>Test set</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfectly bonded</td>
<td>1000</td>
<td>400</td>
<td>600</td>
<td>2000</td>
</tr>
<tr>
<td>Debonded</td>
<td>1200</td>
<td>400</td>
<td>600</td>
<td>2200</td>
</tr>
</tbody>
</table>

Figure 13a,b present the training and validation accuracy trends for epoxy A and epoxy B materials, respectively. Both two curves (Figure 13a,b) demonstrate an increasing tendency of training and validation accuracy with epoch iterations. The trend of the graphs does not show significant oscillation of accuracy over epochs and neither graph does not show significant overfitting or underfitting signs. Then, by applying the trained model on the testing subsets, the confusion matrix was plotted for the epoxy A and epoxy B coating material as shown in Figure 13c,d, respectively.

In this study, the accuracy of the test subset in the confusion matrix is defined by Equation (8):

\[
\text{Accuracy} = \frac{\text{CC}}{\text{CC} + \text{IC}} \times 100\% \tag{8}
\]

where CC indicates correct classified values and IC is incorrect classified values. As evident from the confusion matrix shown in Figure 13b,c, the proposed method provides high accuracy for both types of coating materials. For example, the accuracy of the correct classification of the debonded coating material was 99.67% for the epoxy A type coating material, whereas for epoxy B it was 99.50%. Similarly, correct detection of the bonded section of the coating material was also high for both types of coating materials: 99.83% for epoxy A and 99.33% for epoxy B. However, there was a minor error rate of 0.33% by classifying the bonded section as debonded for epoxy A and 0.50% for epoxy B.

As shown in Figures 8b and 10b, echoes A2 and A3 are reflected from the back wall of the base material to the coating layer. In the spectrogram of the waveform from the bonded layer, echoes A2 and A3 appear, although they do not appear in the debonded coating layer (in Figures 9b and 11b). This feature allows the CNN to classify the bonded and debonded layer with high accuracy. According to the confusion matrix, it can be seen that the CNN model was classifying with over 99% accuracy of bonded and debonded layer for both types of coating materials.
Figure 13. Performance curves of the training and validation subsets for (a) epoxy A and (b) epoxy B. Confusion matrix of the performance test subsets of the CNN for (c) epoxy A and (d) epoxy B.

6. Conclusions

In this study, the STFT method was presented to measure the thickness and bonding state of the coating layer. We also trained a CNN with STFT spectrograms and automatically detected the bonding state of the coating layers using the trained model. Based on the results, the following conclusions can be drawn:

1. The delay line allows for the measuring of the thickness of the coating layer with a single echo from the back wall of the coating layer. The magnitude projection of the STFT allowed for a more accurate measurement of the thickness of the coating material. When comparing the ultrasonic measurement result with the thickness value measured with a caliper, similar results were found and the difference between the two values was about 1%.

2. Although the TOF of the reflected echo can also be evaluated on time–domain waveforms, the advantage of the STFT-based approach concerns the fact that it can accurately and quickly estimate the TOF of a signal even at low signal-to-noise ratios.

3. The ratio of STFT magnitude peaks between two sequential echoes A1 and A0 show a clear difference between the bonded and debonded coatings. The ratio of the STFT size peaks was larger in the case of debonding than in the case of the bonded coating layer. It was also established that the debonded coating layer can be confirmed regardless of the coating material.

4. It is possible and effective to detect the debonded coating layer based on the spectrogram of the waveform and the CNN. The applied CNN-based approach has been shown to accurately classify the bonded and debonded states of coating layers with greater than 99% accuracy. Based on this study, the thickness and bonding state of the coating layer can be easily determined through a combination of the spectrogram of the waveform and a CNN. The proposed method can be quickly implemented on
other types of coating materials. Further optimization of the design parameters will be performed in future studies.


**Funding:** The authors thank the support provided by the National Research Foundation of Korea (NRF) grant as funded by the Korean government (MSIP) (No. 2020M2D2A1A02069933) and the Korea Evaluation Institute of Industrial Technology (KEIT) grant as funded by the Korean government (MOTIE) (No. 10085576). This research was funded by the Korea Institute of Ocean Science and Technology (KIOST), KIOST project number PE99931.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data is contained within the article.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Kim, S.; Hong, H.; Han, T.H.; Kim, M.O. Early-age tensile bond characteristics of epoxy coatings for underwater applications. *Coatings* 2019, 9, 757. [CrossRef]
3. Aurrekoetxea, M.; López de Lacalle, L.N.; Llanos, I. Machining stresses and initial geometry on bulk residual stresses characterization by on-machine layer removal. *Materials* 2020, 13, 1445. [CrossRef]
11. Greenwood, M.S.; Adamson, J.D.; Bamberger, J.A. Long-path measurements of ultrasonic attenuation and velocity for very dilute slurries and liquids and detection of contaminants. *Ultrasonics* 2006, 44, e461–e466. [CrossRef]
17. Li, P.; Chen, S.; Cai, Y.; Chen, J.; Li, J. Accurate TOF measurement of ultrasonic signal echo from the liquid level based on a 2-D image processing method. *Neurocomputing* 2016, 175, 47–54. [CrossRef]


37. Lin, L.; Zhang, W.; Ma, Z.; Lei, M. Porosity estimation of abradable seal coating with an optimized support vector regression model based on multi-scale ultrasonic attenuation coefficient. *NDT E Int.* **2020**, *113*, 102272. [CrossRef]


