applied sciences

MDPI

Article

Denoising of Raman Spectra Using a Neural Network Based on Variational Mode Decomposition, Empirical Wavelet Transform, and Encoder-Bidirectional Long Short-Term Memory

Xuyi Zhang [1,†], Yang Bai [1,2,*,†], Yuan Ma [3], Peidong He [2], Yinhui Tang [2] and Xiaoning Lv [4]

[1] Department of Criminal Science and Technology, Henan Police College, Zhengzhou 450046, China; zhangxuyi@hnp.edu.cn
[2] Aerospace Information Research Institute, Chinese Academy of Sciences, No. 9 Dengzhuang South Road, Haidian District, Beijing 100094, China; hepeidong@aircas.ac.cn (P.H.); tangyinhui18@mails.ucas.ac.cn (Y.T.)
[3] School of Mechanical Engineering, Tsinghua University, Haidian District, Beijing 100084, China; yuanma@tsinghua.edu.cn
[4] Institute of Software, Chinese Academy of Sciences, No. 4 Nan Si Street, Haidian District, Beijing 100089, China; xiaoning@iscas.ac.cn
* Correspondence: baiyang@hnp.edu.cn
† These authors contributed equally to this work.

Abstract: Raman spectroscopy has been widely applied in numerous fields including bioanalysis, disease diagnosis, and molecular recognition, owing to its unique advantages of being non-invasive, rapid, and highly specific. However, the acquisition of Raman spectral data is often susceptible to various noise interferences, such as shot noise from the internal detector and dark current noise from the instrument. As a result, the weak Raman signals typically become arduous to discern, which affects the localization and identification of characteristic spectral peaks. This study investigates variational mode decomposition (VMD), empirical wavelet transform (EWT), and integrated encoder-bidirectional long short-term memory (EBiLSTM) modules to propose a neural network algorithm for adaptive denoising of Raman spectra. By combining VMD and EWT, the Raman spectra are decomposed into several sub-sequences, which solves the problem of mode mixing between high-frequency signals and noise in empirical mode decomposition, and significantly reduces the complexity of the original Raman spectral lines. The correlation coefficient between each modal component and the original signal is calculated along with the zero-crossing rate index to categorize the noise and signal sequences. Leveraging the linear differences between the ideal spectral lines and the noisy spectral curves, an encoder-bidirectional long short-term memory (EBiLSTM) denoising network is constructed for hierarchical denoising to extract valid spectral feature information from the high-frequency components, realizing refined adaptive denoising of the Raman spectra. Industry standard objective evaluation metrics on the signal-to-noise ratio and root mean square error are utilized to conduct simulation experiments comparing state-of-the-art algorithms, including empirical mode decomposition, VMD, sliding window averaging, and wavelet thresholding. The experimental results demonstrate that the Raman spectral denoising algorithm combining variational mode decomposition and neural networks improves the denoising performance by 13.38% to 72%, exhibiting higher accuracy and reliability.

Keywords: variational mode decomposition; empirical wavelet transform; Raman spectroscopy; hierarchical denoising

check for updates

Citation: Zhang, X.; Bai, Y.; Ma, Y.; He, P.; Tang, Y.; Lv, X. Denoising of Raman Spectra Using a Neural Network Based on Variational Mode Decomposition, Empirical Wavelet Transform, and Encoder-Bidirectional Long Short-Term Memory. *Appl. Sci.* **2023**, *13*, 12046. https://doi.org/10.3390/app132112046

Academic Editor: Thomas Lindner

Received: 19 October 2023
Revised: 30 October 2023
Accepted: 1 November 2023
Published: 5 November 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

As a typical optical detection technique, Raman spectroscopy has been widely applied across numerous fields, including bioanalysis, disease diagnosis, and molecular recognition [1–4]. However, the acquisition of Raman spectral data is often subject to various types

of noise interference, such as CCD detector shot noise and dark current noise within the instrument, as well as the influence of the fluorescence background from the material. Especially under the stringent requirements for the on-site detection of special samples, such as a short exposure time and low power excitation lasers, the characteristic information carried in the Raman spectra can even be overwhelmed by the abundant noise. The weak Raman spectral signals usually become difficult to discern, thereby affecting the localization and identification of characteristic spectral peaks [5,6]. Researchers have made abundant efforts and attempts in recent years to minimize the influence of noise, baseline drift, and other interfering factors on the analysis of Raman spectral data during subsequent data processing [7]. Existing Raman spectral denoising methods include fast Fourier transform, wavelet transform, and polynomial fitting, which can effectively eliminate noise in Raman spectra under specific conditions. However, there exist issues related to the empirical selection of filtering parameters, where different cut-off frequencies, filter window lengths, basis functions, thresholds, decomposition levels, and fitting orders can greatly influence the denoising effects [8–10]. Liu et al. [11] employed the Savitzky–Golay convolutional smoothing (SG smoothing) method for further optimization processing of citrus leaf chlorophyll content Raman spectra after background interference subtraction, providing a simple analytical approach for quantitative analysis of citrus leaf chlorophyll content. However, it can cause issues of over smoothing the signal peaks. To address the current challenge of unsatisfactory denoising effects on confocal Raman microscopic images, Fang et al. [12] proposed an omnidirectional denoising algorithm for Raman microscopic images based on wavelet analysis. It first performs wavelet transform on the Raman spectral images, followed by bilateral filtering of the Raman images, thereby suppressing noise interference in the Raman spectral microscopic images, and effectively improving the signal-to-noise ratio and accuracy of compositional analysis. However, when scanning new samples, suitable parameters need to be empirically selected based on human experience to achieve ideal outcomes, lacking flexibility and adaptivity. Li et al. [13] decomposed Raman spectra using empirical mode decomposition and realized adaptive denoising of Raman spectra through low-pass thresholding. However, for biological Raman spectra, the EMD method suffers from mode mixing issues, while low-pass thresholding fails to differentiate between high-amplitude noise and high-frequency Raman signals. Konstantin Dragomiretskiy, Dominique Zosso et al. [14] proposed the variational mode decomposition (VMD) algorithm, which can adaptively decompose a signal into a finite number of band-limited intrinsic mode functions with their own center frequencies. Compared to other methods, the VMD algorithm has more accurate and stable feature extraction capabilities for the low-frequency band of signals. Since the primary features of most signals are contained in the low frequencies, this filtering approach can perform denoising of noisy signals more effectively. Zhang et al. [15] proposed a cherry quality inspection model based on VMD–SVD–MSR, which utilizes VMD to decompose near-infrared reflectance spectra into multiple IMF components and realizes multimodal layer decomposition. The correlation coefficient between each IMF component signal and SSC, as well as the moisture content is calculated separately, with larger correlation coefficients indicating that the corresponding IMF layers are more suitable as feature signals. Subsequently, Si-PLS wavelength selection is applied to further extract optimal spectral bands of the IMF layers, and singular entropy is obtained through singular value decomposition to establish a multivariate stepwise regression prediction model. The above processing based on VMD targets near-infrared spectra, rather than Raman spectra. Near-infrared spectra have smooth characteristic bands and less noise interference, with valid feature information concentrated in the low frequencies. Thus, near-infrared spectral denoising should focus on removing high-frequency noise. In contrast, Raman spectra have more useful information in high or even extremely high frequencies due to their sharp characteristic peaks. Raman spectral denoising should pay more attention to separating feature information from noise in the high frequencies. The high-frequency characteristic information of Raman spectra is often mixed with high-frequency noise. Operating on a low-frequency IMF can easily lead to the loss of high-frequency charac-

teristic information and signal distortion, failing to obtain favorable experimental results. With the development of intelligent optimization algorithms, Zhang et al. [16] searched for the optimal solution for the two parameters of VMD using the fruit fly optimization algorithm (FOA) and selected a corresponding fitness function as the optimization indicator. However, genetic optimization algorithms are relatively complex, with slower iteration and solution processes.

In summary, some common issues in Raman spectral denoising are listed as follows:

- Mode mixing problem: The use of empirical mode decomposition results in mode mixing problems, where high-amplitude noise cannot be differentiated from high-frequency Raman signals. The variational mode decomposition method can resolve the issue of mode mixing in Raman spectral lines, but the residual sequence after VMD decomposition still has complex high-frequency variations that cannot be effectively utilized.
- Complicated and low efficiency: The use of swarm intelligence algorithms, such as genetic algorithms and fruit fly optimization, relies on complex iterative optimization algorithms for parameter tuning, which is time consuming due to the extensive iterations for optimization and the algorithms are difficult to construct.
- Lack of intelligent neural network algorithms: The goal of Raman spectral line denoising is to improve the quality of Raman spectral curves. Therefore, the noise residual between the original signal and the reconstructed spectral curve contains high-amplitude noise and high-frequency Raman signals, which can lead to bias errors in noise estimation. Currently, there are scarce practical applications of neural network models for Raman spectral denoising.

To address the aforementioned issues, this study investigates variational mode decomposition (VMD), empirical wavelet transform (EWT), and integrated encoder-bidirectional long short-term memory (EBiLSTM) modules to propose a neural network algorithm for adaptive denoising of Raman spectra. This enhances the denoising capabilities, while greatly simplifying the complex optimization of the parameters, making it more suitable for automated implementation.

In summary, the main contributions from this work are:

- To address the mode mixing problem in EMD decomposition and the complexity of decomposing high-frequency features in VMD, VMD is utilized to decompose the original load signal into several sub-sequences and a residual sequence, resolving the signal mode mixing issue. EWT is then leveraged to further decompose the residual sequence, significantly reducing the complexity of the original Raman spectral lines. The correlation coefficient between each modal component and the original signal is computed, along with the zero-crossing rate index, to categorize the noise and signal sequences.
- Capitalizing on the linear differences between the ideal spectral lines and noisy spectral curves, a bidirectional LSTM denoising network is constructed for hierarchical denoising, extracting valid spectral feature information from the high-frequency components to accomplish refined adaptive denoising of the Raman spectra.

2. Materials

2.1. Variational Mode Decomposition

VMD can decompose a complex signal into a series of sub-signals with specific bandwidths in the frequency domain, thereby excavating the features of complex nonlinear signals [17]. Assuming the frequency components of each modal component obtained from the VMD decomposition process are concentrated around the center frequency w_k, the signal decomposition problem using VMD can be transformed into an optimization problem. The goal of the optimization is to find the optimal number of modal decompositions K that minimizes the total bandwidth of the decomposed modal sequences, while satisfying the

constraint that the sum of all the decomposed sequences recovers the original signal $f(t)$. The mathematical modeling of this optimization problem can be expressed as:

$$\begin{aligned} &\min_{u_k, w_k}\{\sum_{k=1}^{K}||\partial_t[(\delta(t)+\tfrac{j}{\pi t})*u_k(t)]\mathrm{e}^{-jw_k t}||_2^2\} \\ &\text{s.t.}\sum_{k=1}^{K}u_k = f(t) \end{aligned} \tag{1}$$

In Equation (1): $f(t)$ represents the original load signal; $\delta(t)$ is the Fermi–Dirac distribution function; t is time; u_k and w_k are the modal component and corresponding center frequency for the kth decomposition; K is the number of modal decompositions; j represents the imaginary unit. By introducing a quadratic penalty and Lagrange multipliers into Equation (1), the above constrained variational problem can be transformed into an unconstrained optimization problem, as follows:

$$\begin{aligned} L(u_k, w_k, \lambda) \quad &= \alpha\sum_{k=1}^{K}||\partial_t[(\delta(t)+\tfrac{j}{\pi t})*u_k(t)]e^{-jw_k t}||_2^2 \\ &+||f(t)-\sum_{k=1}^{K}u_k||_2^2+<\lambda(t), f(t)-\sum_{k=1}^{K}u_k> \end{aligned} \tag{2}$$

In Equation (2): $\lambda(t)$ represents the Lagrangian multipliers, α is the balancing parameter of the data-fidelity constraint, $||f(t)-\sum_{k=1}^{K}u_k||_2^2$ is the quadratic penalty term to accelerate the convergence rate. The optimization problem can be solved by finding the minimum of the augmented Lagrangian equation, using the alternating direction method of multipliers.

2.2. Empirical Wavelet Transform

The main steps to further decompose the residual sequence using empirical wavelet transform (EWT) [18] are as follows:

(1) Segment the Fourier spectrum of the residual sequence into N continuous segments;
(2) Construct bandpass filters based on empirical wavelets, where the empirical scaling function $\varphi_n(w)$ and wavelet function $\psi_n(w)$ for each segment are defined as:

$$\varphi_n(w) = \begin{cases} 1, |w| \leq (1-\zeta)w_n \\ \cos[\frac{\pi}{2}\beta(\frac{1}{2\zeta w_n}(|w|-(1-\zeta)w_n))], \\ \qquad (1-\zeta)w_n \leq |w| \leq (1+\zeta)w_n \\ 0, otherwise \end{cases} \tag{3}$$

$$\psi_n(w) = \begin{cases} 1, (1+\zeta)w_n \leq |w| \leq (1-\zeta)w_{n+1} \\ \cos[\frac{\pi}{2}\beta(\frac{1}{2\zeta w_{n+1}}(|w|-(1-\zeta)w_{n+1}))], \\ \qquad (1-\zeta)w_{n+1} \leq |w| \leq (1+\zeta)w_{n+1} \\ \sin[\frac{\pi}{2}\beta(\frac{1}{2\zeta w_n}(|w|-(1-\zeta)w_n))], \\ \qquad (1-\zeta)w_n \leq |w| \leq (1+\zeta)w_n \\ 0, otherwise \end{cases} \tag{4}$$

In the equation: ζ represents the translation factor. The function $\beta(x)$ is defined as [19]:

$$\beta(x) = \begin{cases} 0, x \leq 0 \\ x^4(35-84x+70x^2-20x^3), 0 < x < 1 \\ 1, x \geq 1 \end{cases} \tag{5}$$

(3) Calculate the approximation coefficients $W_f(0,t)$ and detail coefficients $W_f(n,t)$, where the approximation coefficients are obtained by taking the inner product of the

original signal with the empirical scaling functions, and the detail coefficients are obtained by taking the inner product of the original signal with the wavelet functions, defined as:

$$W_f(0,t) = \int f(\tau)\varphi_1(\tau - t)d\tau \tag{6}$$

$$W_f(n,t) = \int f(\tau)\psi_n(\tau - t)d\tau \tag{7}$$

(4) The sub-signals (empirical modes) are calculated, where the first sub-signal $f_1(t)$ and the *nth* sub-signal $f_n(t)$ are defined as:

$$f_1(t) = W_f(0,t) \times \varphi_1(t) \tag{8}$$

$$f_n(t) = W_f(n,t) \times \psi_n(t) \tag{9}$$

2.3. Encoder-Bidirectional Long Short-Term Memory

To address the problem of Raman spectral denoising, an EBiLSTM architecture with a temporal dependency capture module is proposed to consider the contextual sequence information, extract features from the modal components, and determine the types of noise and signal. This enhances the denoising capabilities, while greatly simplifying the complex optimization of the parameters, making it more suitable for automated implementation.

The EBiLSTM structure receives the original Raman spectral signal and continuously learns the multi-dimensional short-term, medium-term, and long-term temporal information and feature information of the spectral data, through the integrated bidirectional input, forget, and output gate structures and memory units [20]. Meanwhile, to address potential overfitting issues during model training, dropout layers [21] and batch normalization layers [22] are designed to reasonably reduce overfitting and improve model robustness. The specific EBiLSTM architecture is illustrated in Figure 1.

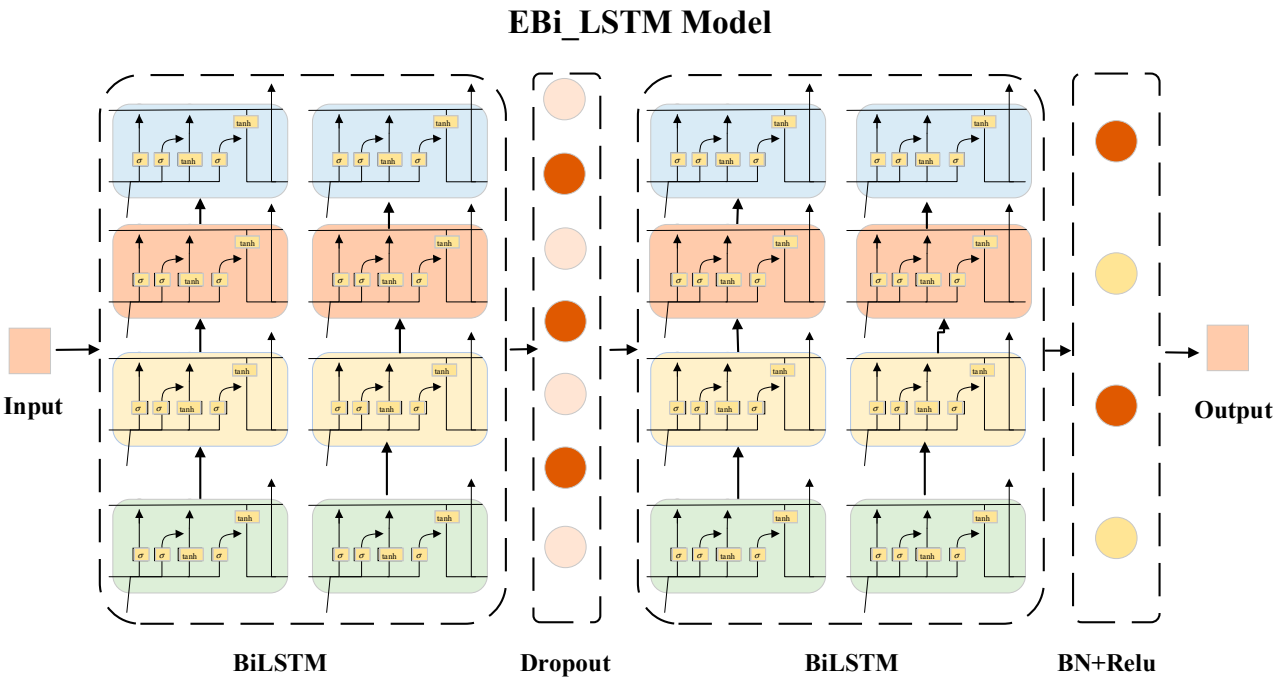


Figure 1. EBiLSTM network architecture.

It can be seen that the EBiLSTM structure contains two BiLSTM layers to capture the long-term dependencies in the sample data, through short-term and long-term neural structures. Each layer uses 32 short-term and long-term neural structures, which are coupled through gate operations, forget layers, learning layers, output layers, and other

structures to output the Raman shift feature information based on the cell state. The weights $W_0^1, W_1^2, W_2^3, \ldots, W_n^m$ are initialized (where n is the total number of hidden nodes and m is the total number of neural structures). Through two consecutive BiLSTM layers, the feature information is captured and processed via dropout for overfitting. Finally, two additional BiLSTM layers are added to incorporate context from the time series information, further integrating the extracted time series information from the shallow network layers. This effectively captures the spectral features from different time series, validly associating short-term, medium-term, and long-term correlations between data of different dimensions, thereby distinguishing the spectral signals from noise.

3. The Proposed Method

3.1. Data Decomposition and Parameter Setting

Before decomposing the original Raman spectral sequence using VMD, the penalty factor α and the number of modal decompositions K need to be pre-determined. They decide the bandwidth size of each modal component. When α increases, the bandwidth of each modal component decreases, and vice versa. Similarly, the value of K directly affects the decomposition results. When K is small, the original signal cannot be fully decomposed, resulting in large decomposition errors and low prediction accuracy. When K is too large, the original signal is over decomposed. The accumulation of errors from excessive decompositions leads to decreased prediction accuracy and increased computational burden.

The number of modal decompositions K is adaptively optimized based on the mean squared error loss function [23], which is specifically defined as:

$$E_{de} = \frac{\sum_{i=1}^{T} (f(t) - \hat{f}(t))^2}{T} \tag{10}$$

In Equation (10): $f(t)$ represents the original signal, $\hat{f}(t)$ denotes the sum of the modal signals from each decomposition, T is the total time. Specifically, $\hat{f}(t)$ is mathematically defined as:

$$\hat{f}(t) = \sum_{k=1}^{K} u_k(t) \tag{11}$$

In Equation (11): $u_k(t)$ represents the *kth* modal component. After defining the loss function, the iterative optimization of the VMD decomposition is performed. The smaller the E_{de} is, the more thorough the VMD decomposition.

After VMD decomposition, the residual sequence is further decomposed using EWT to reduce the complexity of the load sequence. Having too few or too many modal decompositions in the EWT will affect the decomposition results. Considering the convergence characteristics of the center frequencies in VMD, the boundaries and center frequencies of the modal components can be adaptively detected. Therefore, before decomposing the residual sequence using ETW, VMD is first utilized to decompose the residual sequence and the number of ETW modal decompositions is adaptively determined based on the convergence of the center frequencies.

3.2. A Raman Spectral Denoising Model Using VMD–EWT–EBiLSTM

To improve the denoising accuracy of Raman spectra, this paper proposes a Raman denoising model based on VMD, EWT, and EBiLSTM modules. The prediction flowchart of the model is shown in Figure 2 below. The signal components decomposed from the original signal by VMD are called intrinsic mode functions, abbreviated as IMF.

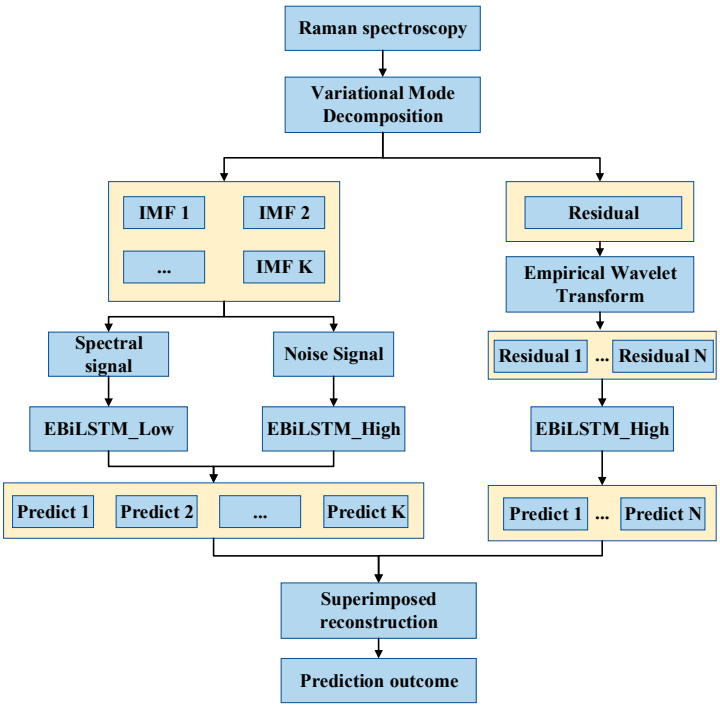


Figure 2. Raman denoising prediction process flowchart.

First, VMD is utilized to decompose the original Raman spectral sequence into several sub-sequences and a residual sequence. EWT is then used to further decompose the residual sequence. The sub-sequences from the VMD decomposition are categorized into spectral signals and noise signals using the zero-crossing rate and correlation coefficient. The zero-crossing rate [24] is specifically defined as:

$$Z_0 = \frac{z_0}{N} \tag{12}$$

In Equation (12): Z_0 denotes the zero-crossing rate, z_0 is the number of zero crossings, N is the total number of load samples.

The correlation coefficient [25] between the modal sequences from the VMD decomposition and the original signal is calculated as specifically defined below:

$$coor = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(v_i - \bar{v})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(v_i - \bar{v})^2}} \tag{13}$$

where: x represents the original signal, v represents the modal sequence.

Next, the categorized spectral signal sub-sequences are input into a network structure with a small number of stacked EBiLSTM modules, termed EBiLSTM_Low, to capture the long-term dependencies of the spectral sequence features and extract the features with low noise interference. The categorized noise sequences are input into a stack of high-frequency EBiLSTM modules, termed EBiLSTM_High, for further feature extraction and noise identification to eliminate additional noise. Finally, the signal and noise identification results for each sub-sequence are superimposed to reconstruct the overall denoised spectral result.

The designed Raman spectral prediction model based on VMD–EWT–EBiLSTM is illustrated in Figure 3. The backbone network for the spectral signal components captures the feature temporal correlations and extracts information through stacked EBiLSTM modules, with 1, 2, 4, 2 layers stacked. The backbone network for the noise signals extracts local Raman spectral features and noise signals through high-frequency stacked IASSP

structures with 2, 2, 2, 4, 4, 2 modules stacked, jointly forming a dual-channel spectral signal and noise feature extraction and denoising model. The extracted feature information from the dual-channel links is superimposed for reconstruction, and input into 64 short-term and long-term neural structures to learn the fused features and noise characteristics from both paths. The features captured by the short-term and long-term neurons are then connected through a skip structure. After concat layer processing, it is input into the fully connected layer, and dropout layers are utilized to avoid overfitting. Finally, spectral signal and interference noise discrimination is performed through a sigmoid activation function.

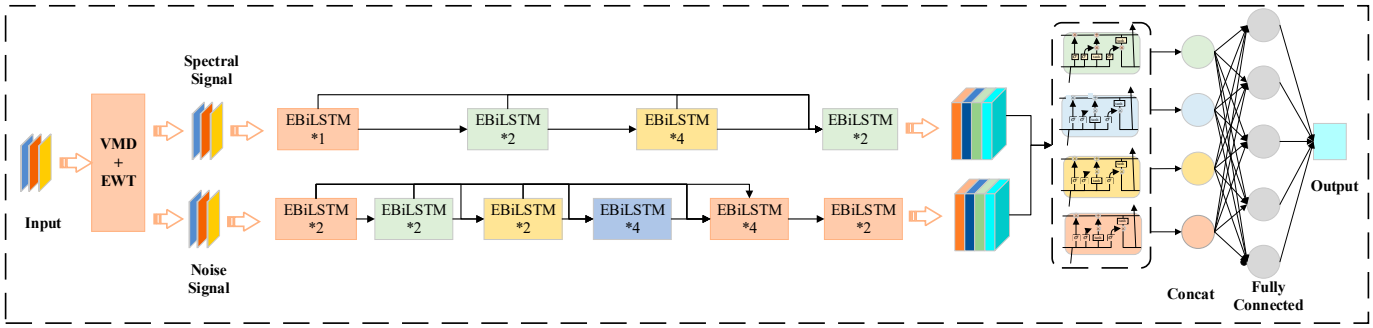


Figure 3. Raman spectra prediction model designed with VMD–EWT–EBiLSTM.

4. Results and Analysis

4.1. Dataset and Raman Spectral Sequence Decomposition

The experiments utilized a Senterra Raman spectrometer to scan plant leaf samples, with a laser power of 0.5 mW, a wavelength of 785 nm, and a spectral scanning range of 252–2909 cm^{-1}. The training dataset was acquired from Raman spectra collected over a period of time, with a length $L = 2000$. The spectral dataset was evenly segmented into $m = 10$ sections, with $n = 200$ data points per segment, giving a total of $mn = L$ points. The originally collected spectral data are denoted as S_j^{pi}, and the noisy datasets with randomly added 20 dB and 30 dB noise are denoted as X_j^{pi}, where p indicates the pth spectrum in the training set, i indicates the segment number, and j indicates the data point within the segment. The noisy data and labels were fed into the neural network for training to obtain the Raman spectral denoising model.

An example segment of the noisy plant leaf spectra is shown in Figure 4 below.

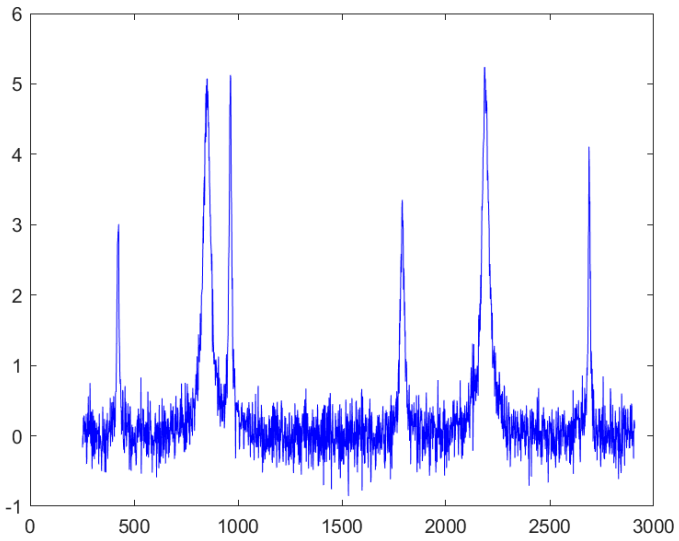


Figure 4. Noisy spectra of plant leaves.

The VMD algorithm decomposes the Raman spectral sequences with initial parameters set as: α value is set to 4000 to adjust the bandwidth of each modal component to a proper level, τ is set to 0.00001, λ is set to 0, K is initialized as 1, and the threshold of E_{de} is set to 0.03. The parameters are then iteratively updated using the VMD decomposition algorithm, calculating the decomposition error E_{de} each time.

After iterative updates to the VMD algorithm decomposing the Raman sequences, the iterations terminate when the number of modal decompositions $K = 5$, with a decomposition error of 0.0267. Therefore, $K = 5$ is taken as the number of variational modal decompositions, and the specific decomposition results are shown in Figure 5. The number of modal decompositions in the EWT is determined by the convergence of the center frequency values when decomposing the VMD residual sequence. Finally, when the number of modes N = 4, the center frequency values converge, so the number of modal decompositions in the EWT is taken as N = 4. The specific decomposition results are shown in Figure 6.

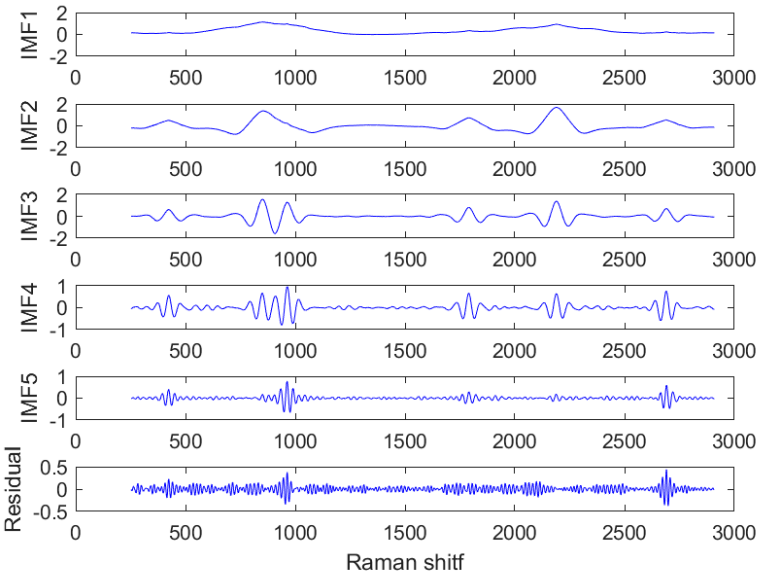


Figure 5. VMD decomposes the original signal.

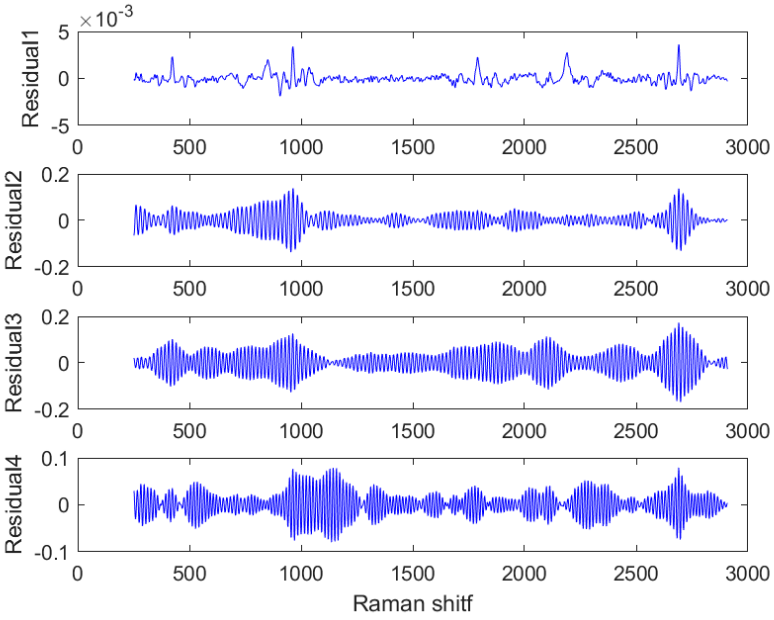


Figure 6. EWT breaks down the residual signal.

The decomposed sequences are categorized into spectral and noise components using the zero-crossing rate and correlation coefficient. The zero-crossing rate of each modal component are calculated as shown in Table 1. The modal components with zero-crossing rates greater than 0.05 and the correlation coefficients less than 0.5 are taken as noise sequences, and the remaining are taken as spectral energy components.

Table 1. Zero-crossing rate of different modal components.

Intrinsic Mode Functions	IMF1	IMF2	IMF3	IMF4	IMF5
Z_0	0	0.0080	0.0268	0.065	0.498

4.2. Loss Function and Evaluation Metrics

During network training, in order to obtain optimal Raman spectral prediction results, there are two stages. In the first stage, only the backbone noise signal component feature and spectral temporal dependency capture branches are trained. In the second stage, the overall network structure is trained based on the weights from the first stage training. The Adam optimization algorithm [26] is utilized to find the global optimum during training. The loss function of the neural network prediction model is defined as:

$$L_\gamma(y, y^P) = \sum_{i:y_i<y_i^p} (1-\gamma)\left|y_i - y_i^p\right| + \sum_{i:y_i>=y_i^p} \gamma\left|y_i - y_i^p\right| \tag{14}$$

In Equation (14): γ is the desired quantile, between 0 and 1, y represents the original signal value of the model, referring to the signal before the addition of noise, y_i^p is the predicted value of the model. By minimizing the loss function, the parameters of the neural network model are iteratively updated.

The mean squared error (MSE) [27] and signal-to-noise ratio (SNR) [28] are utilized, as two common objective evaluation metrics, to comparatively analyze the performance of different denoising algorithms.

The mean squared error can measure the deviation between the denoised signal and the original signal, effectively reflecting the denoising accuracy. It is defined as:

$$MSE = \sum_{i=1}^{L} \frac{(X(i) - Y(i))^2}{L} \tag{15}$$

where: $X(i)$ represents the original signal, which refers to the signal before the addition of noise, $Y(i)$ represents the denoised signal.

The signal-to-noise ratio is the ratio of signal power to noise power, which can objectively indicate the presence of noise in the signal. It is defined as:

$$SNR = 10\log_{10}(\frac{E_{signal}}{E_{noise}}) \tag{16}$$

$$E_{signal} = \sum_{n=1}^{L} (S_{ori}(n) - \bar{S_{ori}})^2 \tag{17}$$

$$E_{noise} = \sum_{n=1}^{L} (S_{ori}(n) - S'(n))^2 \tag{18}$$

where: E_{signal} represents the signal energy, E_{noise} represents the noise energy, S_{ori} is the ideal spectrum, S' is the denoised Raman spectral signal, $\bar{S_{ori}}$ is the mean of the original spectral signal, n is the Raman shift, L is the length of the Raman spectrum.

4.3. Experimental Results from Comparative Study on Different Decomposition Algorithms

To verify the effectiveness of the VMD–EWT algorithm in decomposing Raman spectral sequences for spectral denoising, comparative experiments are conducted using the EMD [29], VMD [30], and EBiLSTM network only, and the VMD–EWT–EBiLSTM network structure. The Raman spectral curves from the test set are randomly selected for validation, with the results shown in Figures 7 and 8.

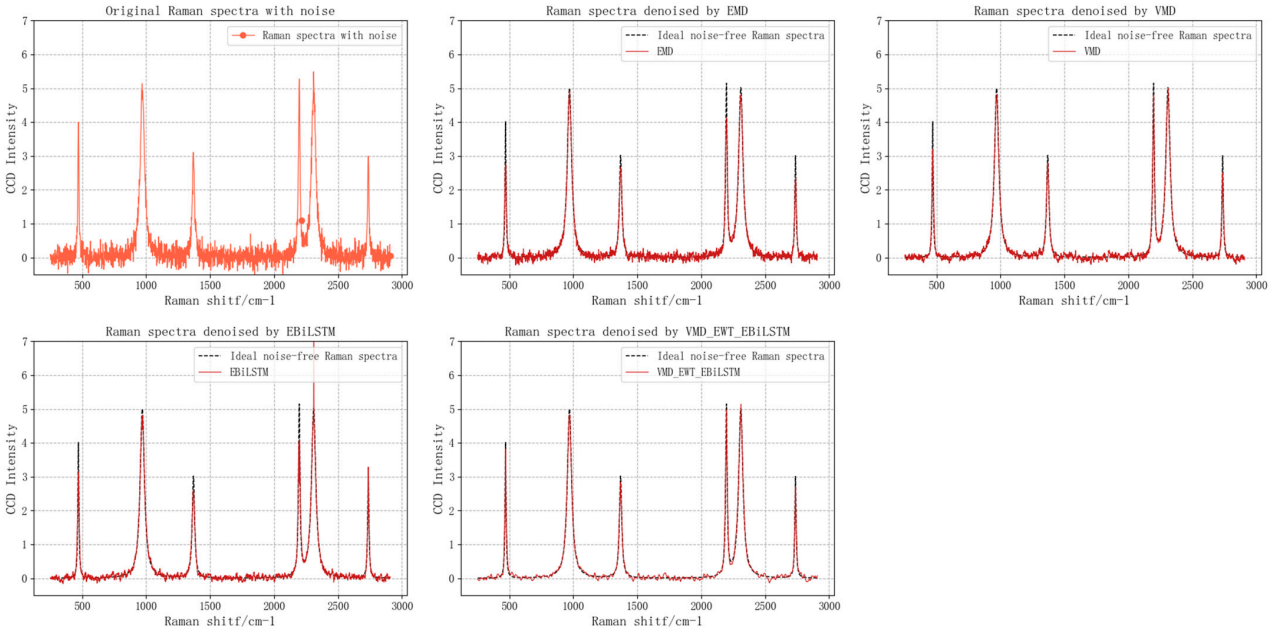


Figure 7. Comparative study on different decomposition denoising algorithms for 20 dB noisy signals.

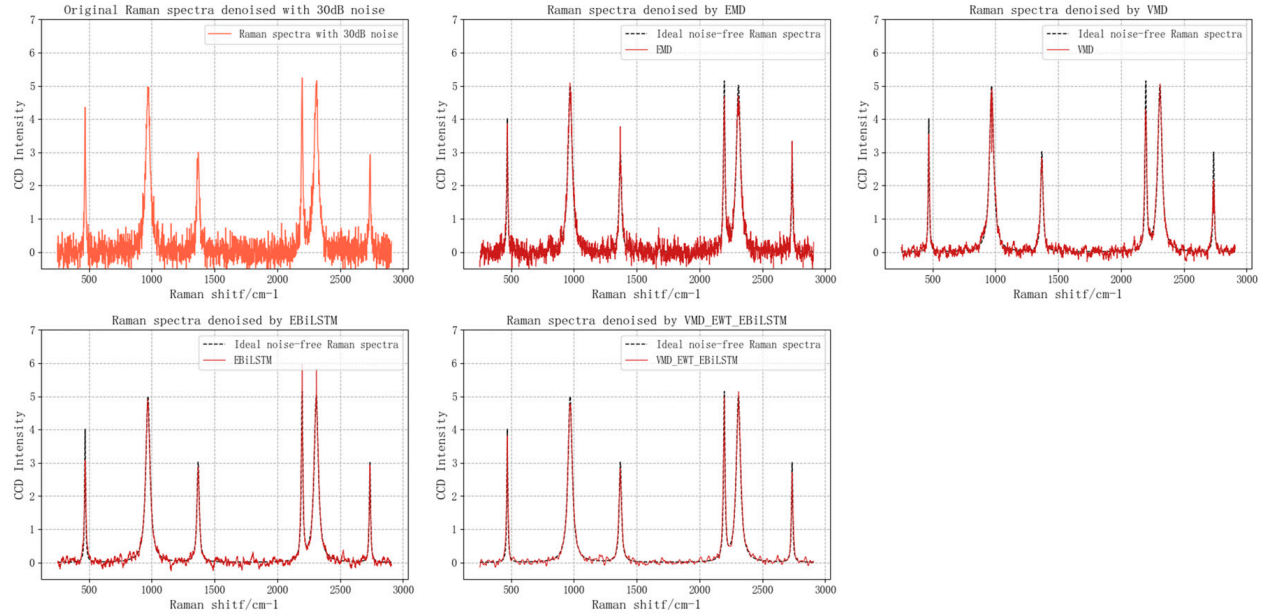


Figure 8. Comparative study on different decomposition denoising algorithms for 30 dB noisy signals.

Gaussian noise at 20 dB and 30 dB is added to the collected Raman spectral sequences. Under low noise interference, all denoising methods achieve the Raman spectral denoising task relatively well. However, with increasing noise, the denoising performance by the EMD and VMD methods decreases markedly. Using only the EBiLSTM method leads

to larger fluctuations in denoising and some errors compared to the ideal spectral curve around the peaks. The VMD–EWT–EBiLSTM network structure used in this paper provides denoising results closer to the true values, and better denoising performance.

Objective metrics on the signal-to-noise ratio (SNR) and mean squared error (MSE) are further utilized to evaluate the performance of different modal decomposition methods. The specific results are shown in Table 2 below. As shown in Table 2, the modal decomposition method proposed in this paper achieves a lower MSE and a higher SNR, with the MSE reduced by 23–56% and the SNR improved by 10–66% during denoising. This further verifies that under interference from different noise levels, using the VMD–EWT decomposition method to extract feature components helps improve the denoising accuracy for the Raman model.

Table 2. Comparison of various decomposition methods for MSE and SNR.

		Raman Spectral Sequence with Noise	EMD	VMD	EBiLSTM	VMD–EWT–EBiLSTM
20 dB	MSE	14.5969	5.7591	3.8723	3.2742	2.5137
	SNR	11.7945	15.7468	20.3494	23.6497	26.1491
30 dB	MSE	24.1763	9.4766	5.6714	3.7818	2.9173
	SNR	4.6342	10.3823	14.1918	18.6463	24.7691

4.4. Experimental Results Comparing Different Raman Spectral Denoising Algorithms

To further validate the performance of the VMD–EWT–EBiLSTM network architecture for Raman spectral denoising, comparative experiments are conducted with four state-of-the-art denoising methods, including Fourier transform [31], Savitzky–Golay filtering [11], sliding window averaging [32], and wavelet thresholding [33].

A Gaussian noise of 20 dB and 30 dB is added to the collected Raman spectral sequences. The prediction results are shown in Figures 9 and 10. It can be seen that for the Raman spectra with 20 dB of Gaussian noise, the denoising results for the VMD–EWT–EBiLSTM, wavelet thresholding, and sliding window averaging are consistent with the ideal Raman spectra trend, while Fourier transform and S-G filtering result in little deviation from the ideal values. As the noise increases, processing Raman spectra with 30 dB of Gaussian noise, the denoising effects of sliding window averaging and wavelet thresholding fluctuate more, with some deviation around the peaks compared to the ideal spectrum. Fourier transform and Savitzky–Golay filtering do not provide satisfactory denoising effects. The VMD–EWT–EBiLSTM used in this paper maintains stable denoising performance.

Objective metrics on the signal-to-noise ratio (SNR) and mean squared error (MSE) are further utilized to evaluate the performance of different denoising methods. The specific results are shown in Table 3. It can be intuitively seen that the denoising model proposed in this paper has relatively lower errors. Compared to existing mainstream prediction algorithms, the denoising performance is improved by 13.38% to 72%, better demonstrating the superior denoising effects of the VMD–EWT–EBiLSTM network architecture on Raman spectral curves.

Table 3. Comparison of various denoising methods for MSE and SNR.

		Raman Spectral Sequence with Noise	Fourier Transform	Sliding Window Averaging	Savitzky–Golay	Wavelet Thresholding	VMD–EWT–EBiLSTM
20 dB	MSE	15.1737	8.3794	2.9908	6.3724	2.6177	2.1164
	SNR	8.4739	14.4940	21.8077	18.7647	22.0222	24.9691
30 dB	MSE	29.3914	16.8756	6.1258	13.0745	4.1837	3.2727
	SNR	4.7192	8.3827	13.4691	10.7281	16.6238	20.7194

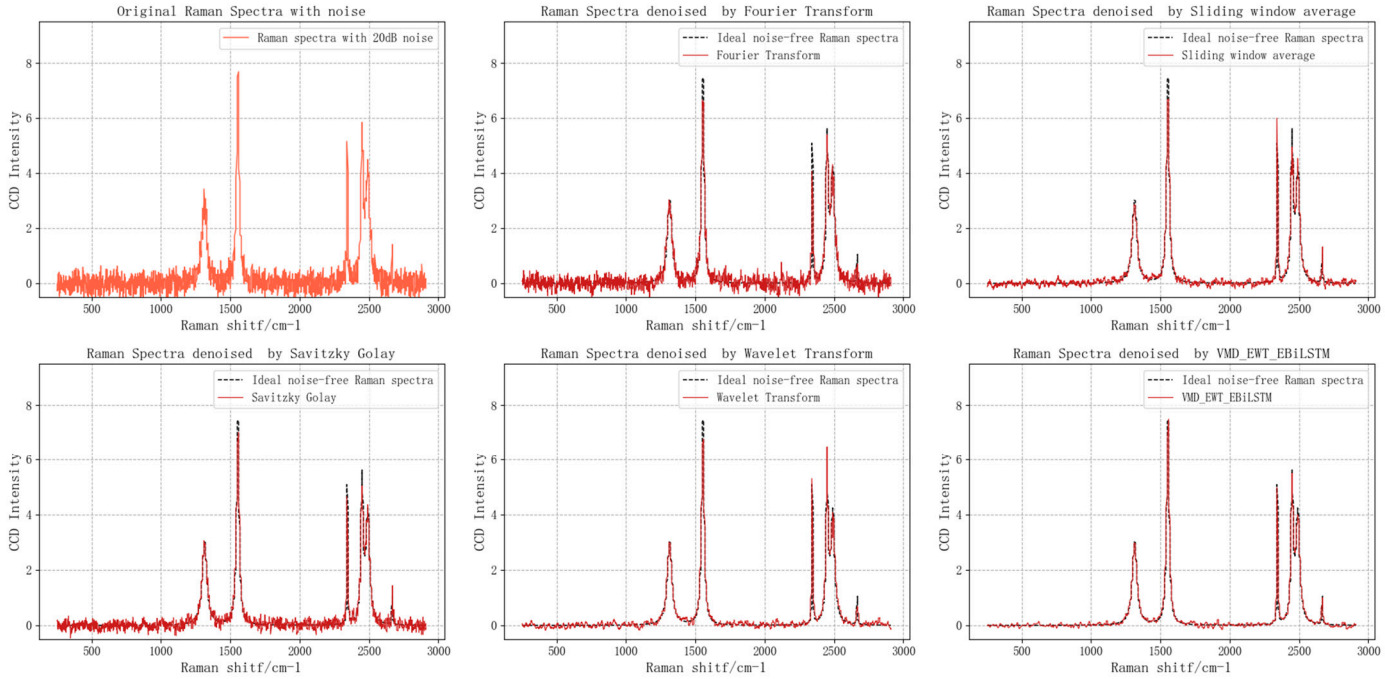


Figure 9. Comparative study on different denoising algorithms for 20 dB noisy signals.

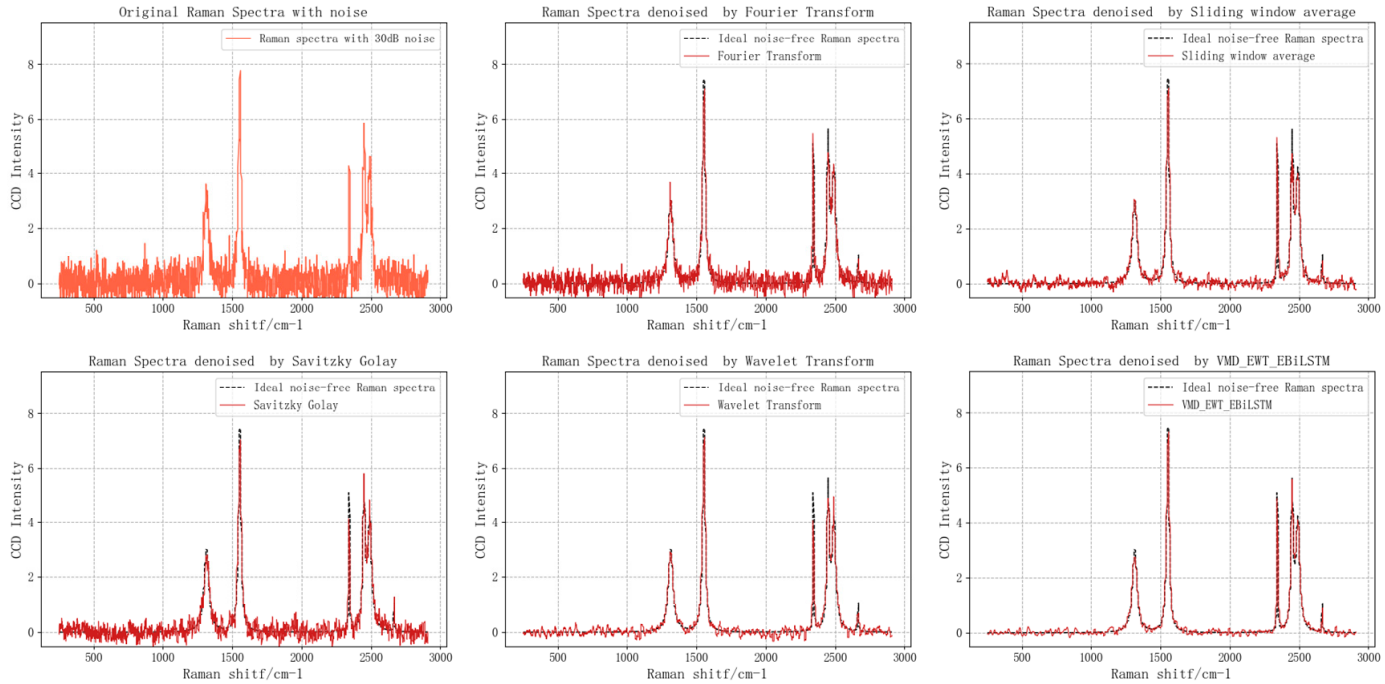


Figure 10. Comparative study on different denoising algorithms for 30 dB noisy signals.

4.5. Raman Denoising Experimental Results

To thoroughly validate the denoising performance of the VMD–EWT–EBiLSTM network architecture, comprehensive performance analysis is carried out on the test set under 20 dB and 30 dB random noise conditions, by examining the denoising outcomes from each algorithm. The MSE and SNR objective metrics are utilized to quantify the results, which are presented in Figures 11 and 12.

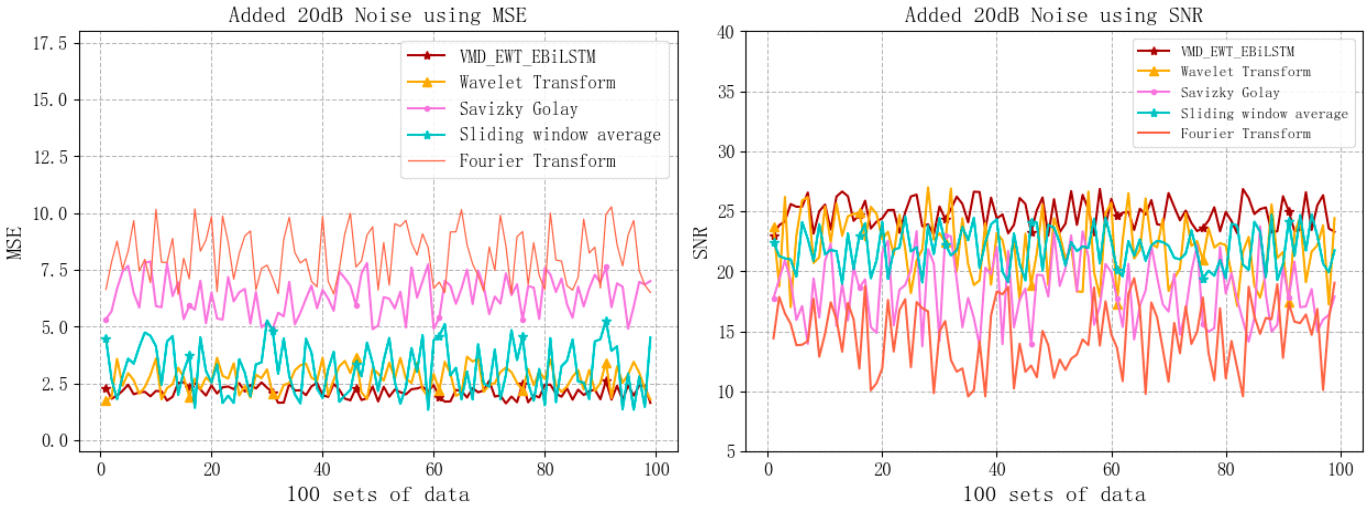


Figure 11. Comparison of MSE and SNR metrics for different denoising methods for a 20 dB noisy signal.

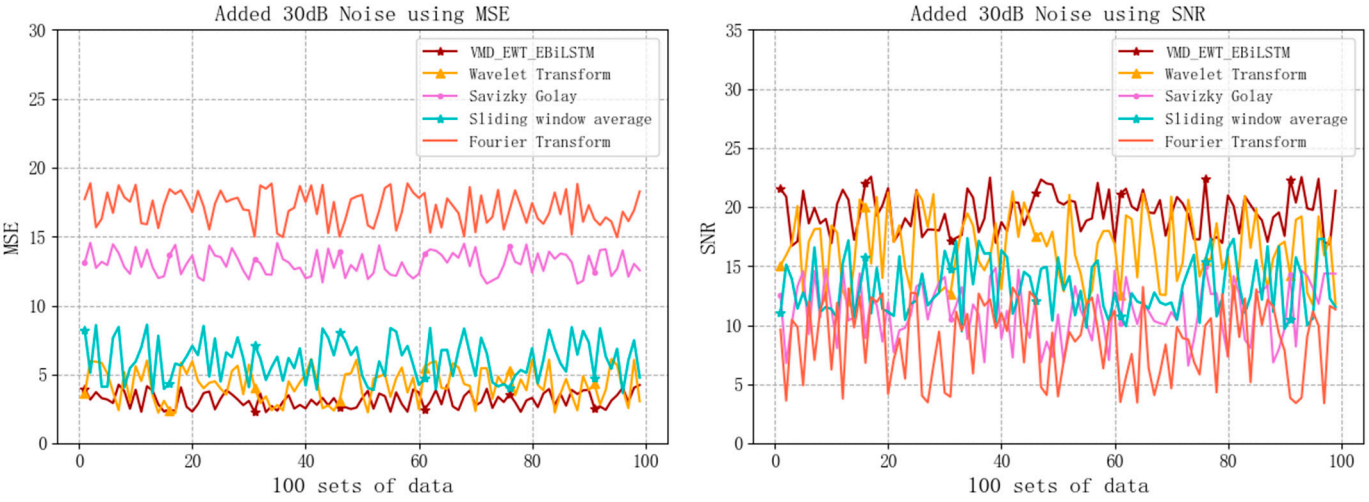


Figure 12. Comparison of MSE and SNR metrics for different denoising methods for a 30 dB noisy signal.

It can be observed from Figures 11 and 12 that for both 20 dB and 30 dB Raman spectra, the VMD–EWT–EBiLSTM network architecture has a lower MSE and a higher SNR compared to other benchmark denoising algorithms, demonstrating superior denoising performance. Table 4 summarizes the average MSE and average SNR over the entire Raman spectral test set. Here, MSE_Mean denotes the average MSE across all the spectral test samples, and SNR_Mean denotes the average SNR across the spectral test set. Based on these metrics, it is evident that the VMD–EWT–EBiLSTM network proposed in this paper provides better denoising results.

Table 4. Comparison of various methods for MSE and SNR for different noise levels.

		Fourier Transform	Sliding Window Averaging	Savitzky–Golay	Wavelet Thresholding	VMD–EWT–EBiLSTM
20 dB	MSE_Mean	8.3024	3.1983	6.4614	2.7001	2.1186
	SNR_Mean	14.7124	21.9335	18.5184	21.8740	24.8642
30 dB	MSE_Mean	16.8328	6.1053	12.9709	4.3261	3.225
	SNR_Mean	8.4663	13.5375	10.3586	16.6832	19.8631

5. Discussion

To verify the effectiveness of the VMD–EWT algorithm in decomposing Raman spectral sequences for spectral denoising, comparative experiments were conducted with different modal decomposition-based spectral denoising algorithms. The results demonstrate that utilizing VMD to decompose the original sequence into several sub-sequences and a residual sequence, followed by further decomposition of the residual sequence using EWT, and categorizing the noise and signal sequences based on the correlation coefficient and zero-crossing rate index, can resolve the signal mode mixing problem and maintain good denoising performance.

In the comparative denoising experiments, different state-of-the-art denoising algorithms are evaluated using MSE and SNR as objective metrics. The proposed neural network model demonstrates superior denoising performance over other methods, validating the efficacy of the EBiLSTM architecture design. The incorporation of two short-term and long-term neural structures captures the correlations and spectral features in the Raman data.

To thoroughly validate the denoising performance of the VMD–EWT–EBiLSTM network architecture, comprehensive performance analysis is carried out on the test set under 20 dB and 30 dB of random noise conditions, by examining the denoising outcomes from each algorithm. It is evident from the results that the VMD–EWT–EBiLSTM network architecture achieves a lower MSE and a higher SNR compared to other benchmark denoising algorithms, demonstrating superior denoising performance. Furthermore, the analysis of the average MSE and average SNR metrics over the Raman spectral test set provides further evidence that the network architecture yields improved denoising results.

The proposed method in this paper does not always perform ideally. As the noise level increases, the MSE and SNR metrics for noise estimation deteriorate, as shown in Table 4. Under 30 dB of Gaussian noise, the denoising performance of the VMD–EWT–EBiLSTM network declines, mainly due to the following reasons: at high noise levels, the noise intensity becomes comparable to the signal, making it difficult to differentiate between the signal and the noise.

The proposed method in this paper adopts a VMD–EWT–EBiLSTM network architecture for denoising Raman spectral signals. Since it involves relatively complex algorithms, the neural network-based denoising process requires a certain amount of training time and computational resources. To reduce the subsequent time and computation costs, future research could focus on model compression and pruning of the structure and parameters, as well as leveraging GPUs to accelerate model training and inference procedures, in order to further improve the denoising performance for Raman spectra.

6. Conclusions

Aiming to address the issue of noise interference, such as shot noise and dark current noise from the instrument, negatively affecting the localization and identification of spectral characteristic peaks, this paper proposes a new Raman spectral denoising method. VMD is utilized to decompose the original sequence into several sub-sequences and a residual sequence, followed by further decomposition of the residual sequence using EWT. The zero-crossing rate and correlation coefficient are used to categorize spectral and noise signals, and an EBiLSTM network architecture is constructed for spectral feature extraction and noise identification. Through different experimental validations, the proposed new method outperforms comparative algorithms.

(1) Using the VMD–EWT decomposition algorithm achieves higher denoising accuracy compared to other modal decomposition methods, with the MSE reduced by 23–56% and the SNR improved by 10–66%.
(2) The constructed neural network Raman spectral denoising model demonstrates 13.38–72% enhanced denoising performance over state-of-the-art neural networks, reflecting the more stable denoising capabilities of the VMD–EWT–EBiLSTM network structure.

(3) Future improvements to the network model should focus on optimizing the computational efficiency and processing speed.

Author Contributions: Conceptualization, Y.B. and X.Z.; methodology, X.Z. and P.H.; software, X.Z., Y.B. and Y.T.; validation, X.Z. and Y.M.; formal analysis, X.Z. and P.H.; investigation, Y.B. and X.Z.; resources, Y.T. and X.L.; data curation, X.L.; writing—original draft preparation, X.Z.; writing—review and editing, Y.B. and X.Z.; visualization, X.Z.; supervision, Y.B. and Y.M.; project administration, X.Z. and Y.B.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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. Cicchi, R.; Cosci, A.; Rossari, S.; Kapsokalyvas, D.; Baria, E.; Maio, V.; Massi, D.; De Giorgi, V.; Pimpinelli, N.; Saverio Pavone, F. Combined fluorescence-Raman spectroscopic setup for the diagnosis of melanocytic lesions. *J. Biophotonics* **2013**, *7*, 86–95. [CrossRef]
2. Hu, C.; Chen, S.; Chen, J.; Zhang, W.; Chen, X. Applications of Raman spectroscopy technology in polymer research. *Polym. Bull.* **2014**, *3*, 30–45.
3. Zhou, M.; Liao, C.; Ren, Z.; Fan, H.; Bai, J. Bioimaging technologies based on surface-enhanced Raman spectroscopy and their applications. *Chin. Opt.* **2013**, *6*, 633–642.
4. Zhai, L.; Zhan, X. Applications of X-ray fluorescence spectroscopy in petrochemical product analysis. *J. Anal. Sci. Technol.* **2015**, *34*, 484–496.
5. Chen, Z.; Zhang, F.; Zhou, Y.; Huan, H. Removal method of multiplicative and additive random noise in spectral signal. *Acta Opt. Sin.* **2017**, *37*, 1–8.
6. Fan, X.; Wang, X.; Wang, X.; Xu, Y.; Que, J.; Wang, X.; He, H.; Li, W.; Zuo, Y. Research on Raman spectral denoising method based on feature extraction for low SNR. *Spectrosc. Spectr. Anal.* **2016**, *36*, 4082–4087.
7. Adi, L.; Einwal; Ai, W. Research progress on Raman spectroscopy. *West Leather* **2017**, *6*, 6–7.
8. Chen, C.; Xu, D.; Cheng, Q. Application of Wavelet Packet Transform for Raman Spectral Data Compression. *J. Yangtzc Univ. Nat. Sci. Ed.* **2008**, *5*, 31–35.
9. Wang, Z.; Liu, M.; Liu, E.; Dong, Z.; Cai, S.; Yin, L.; Liu, F. Raman SNR evaluation method based on extreme value statistics and its applications. *Spectrosc. Spectr. Anal.* **2019**, *39*, 1080–1085.
10. Wei, C.; Wang, J.; Zhang, B.; Dong, Z.; Guan, J. Mineral oil pattern classification based on combinative technology of multi-order derivative Raman spectra. *J. Anal. Sci. Technol.* **2021**, *40*, 747–753.
11. Liu, Y.D.; Cheng, M.J.; Hao, Y.; Zhang, Y.; Hou, Z. Quantitative Analysis of Chlorophyll Content in Citrus Leaves by Raman Spectroscopy. *Spectrosc. Spectr. Anal.* **2019**, *39*, 1768–1772.
12. Fang, S.; Shao, R.; Qiu, L.; Wang, Y. Denoising method for confocal Raman images based on wavelet transform. *Opt. Tech.* **2019**, *45*, 330–335.
13. Li, Q.; Zhang, G.; Liu, Y. Study on Raman spectral denoising method based on EMD. *Spectrosc. Spectr. Anal.* **2009**, *29*, 142–145.
14. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]
15. Zhang, P.; Wang, L.; Ren, L. Cherry quality information detection based on VMD-SVD-MSR model. *J. Tangshan Coll.* **2022**, *3*, 18–24.
16. Zhang, Y.; Zhao, X.; Wu, G.; Jiang, Z.L. Parameter Optimization-Based Bearing Fault Diagnosis Method and System Using FOA-VMD. CN Patent 34562356, 17 August 2022.
17. Yan, X.; Jia, M. Application of CSA-VMD and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings. *Mech. Syst. Signal Process.* **2019**, *122*, 56–86. [CrossRef]
18. Hu, J.; Wang, J. Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. *Energy* **2015**, *93*, 1456–1466. [CrossRef]
19. Gilles, J. Empirical wavelet transform. *IEEE Trans. Signal Process.* **2013**, *61*, 3999–4010. [CrossRef]
20. Huang, Z.; Xu, W.; Yu, K. Bidirectional LSTM-CRF Models for Sequence Tagging. *Comput. Sci.* **2015**, *9*, 1508–1518.
21. Gal, Y.; Ghahramani, Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *JMLR* **2015**, *1*, 142–152.
22. Li, Y.; Wang, N. Adaptive Batch Normalization for practical domain adaptation. *Pattern Recognit. J. Pattern Recognit. Soc.* **2018**, *80*, 18–27. [CrossRef]

23. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [CrossRef]
24. Joo, S.; Choi, J.; Kim, N.; Lee, M.C. Zero-crossing rate method as an efficient tool for combustion instability diagnosis. *Exp. Therm. Fluid Sci.* **2021**, *1*, 123–131. [CrossRef]
25. Ye, J. Fuzzy decision-making method based on the weighted correlation coefficient under intuitionistic fuzzy environment. *Eur. J. Oper. Res.* **2010**, *205*, 202–204. [CrossRef]
26. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2015**, arXiv:1412.6980.
27. Wang, Z.; Wan, D.; Shan, C.; Li, Y.; Zhou, Q. Raman spectral denoising method based on backpropagation neural network. *Spectrosc. Spectr. Anal.* **2022**, *42*, 1553–1560.
28. Zhao, X.; He, Y.; Tong, L. Wavelet Denoising Method for Raman Spectroscopy based on EMD Decomposition. *J. Heilongjiang Bayi Agric. Univ.* **2019**, *31*, 81–86.
29. Zhang, Y.; Zhou, X.; Wang, J.; Guo, Q. Research on ultrasonic signal denoising method based on improved EMD. *Nanjing Univ. Posts Telecommun.* **2016**, *36*, 49–52.
30. Lu, J.; Guo, J.; Qu, X. Research on denoising algorithm based on VMD-OV. *Chem. Ind. Autom. Instrum.* **2021**, *48*, 142–145.
31. Cao, S.; Zheng, J.; Pan, H. Raman spectral denoising method based on improved adaptive empirical Fourier decomposition. *J. Vib. Shock.* **2022**, *41*, 287–299.
32. Jiang, C.; Sun, Q.; Liu, Y.; Liang, J.; An, Y.; Liu, B. A New Peak Detection Algorithm of Raman Spectra. *Spectrosc. Spectr. Anal.* **2014**, *34*, 103–107.
33. Li, H.; Zhang, S. Application of an improved wavelet threshold function in Raman Spectra denoising. *Electron. Des. Eng.* **2022**, *24*, 172–175.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.