Abstract: The micro-electro-mechanical system (MEMS) gyroscope is a micro-mechanical gyroscope with low cost, small volume, and good reliability. The working principle of the MEMS gyroscope, which is achieved through Coriolis, is different from traditional gyroscopes. The MEMS gyroscope has been widely used in the fields of micro-inertia navigation systems, military, automotive, consumer electronics, mobile applications, robots, industrial, medical, and other fields in micro-inertia navigation systems because of its advantages of small volume, good performance, and low price. The material characteristics of the MEMS gyroscope is very significant for its data output, and the temperature determines its accuracy and limits its further application. In order to eliminate the effect of temperature, the MEMS gyroscope needs to be compensated to improve its accuracy. This study proposed an improved variational modal decomposition—extreme learning machine (VMD-ELM) algorithm based on convolutional neural networks—long short-term memory (CNN-LSTM) and particle swarm optimization—support vector machines (PSO-SVM). By establishing a temperature compensation model, the gyro temperature output signal is optimized and reconstructed, and the gyro output signal with better accuracy is obtained. The VMD algorithm separates the gyro output signal and divides the gyro output signal into low-frequency signals, mid-frequency signals, and high-frequency signals according to the different signal frequencies. Once again, the PSO-SVM model is constructed by the mid-frequency temperature signal to find the temperature error. Finally, the signal is reconstructed through the ELM neural network algorithm, and then, the gyro output signal after noise is obtained. Experimental results show that, by using the improved method, the output of the MEMS gyroscope ranging from $-40$ to $60 \degree C$ reduced, and the temperature drift dramatically declined. For example, the factor of quantization noise (Q) reduced from $1.2419 \times 10^{-4}$ to $1.0533 \times 10^{-6}$, the factor of bias instability (B) reduced from $0.0087$ to $1.8772 \times 10^{-4}$, and the factor of random walk of angular velocity (N) reduced from $2.0978 \times 10^{-3}$ to $1.4985 \times 10^{-6}$. Furthermore, the output of the MEMS gyroscope ranging from $60$ to $-40 \degree C$ reduced. The factor of Q reduced from $2.9808 \times 10^{-4}$ to $2.4430 \times 10^{-6}$, the factor of B reduced from $0.0145$ to $7.2426 \times 10^{-4}$, and the factor of N reduced from $4.5072 \times 10^{-5}$ to $1.0523 \times 10^{-5}$. The improved algorithm can be adopted to denoise the output signal of the MEMS gyroscope to improve its accuracy.

Keywords: MEMS gyroscope; convolutional neural networks—long short-term memory (CNN-LSTM); particle swarm optimization—support vector machines (PSO-SVM); variational modal decomposition (VMD); temperature compensation

1. Introduction

The micro-electro-mechanical system (MEMS) gyroscope has become a promising sensing technology to arouse research interests due to its advantages of small size, high accuracy, and long durability, and it is widely used in the fields of inertial navigation and positioning. Meanwhile, it has disadvantages including long initial installation [1].
and calibration time [2], more temperature-sensitive equipment and materials [3], and accumulated angle integration errors over time. Specifically, due to the difference in the expansion coefficient of the internal construction materials of the MEMS gyroscope, corresponding thermal resistance and thermal stress will be generated [4], and the change of the environment temperature has a greater impact on the stability of the MEMS gyroscope’s bias. To address these problems, the temperature drift model is usually established to suppress the temperature drift of the gyroscope and improve the accuracy of the MEMS gyroscope [5,6], the relationship between the output of the gyroscope and the temperature drift is predicted and compensated by analyzing the model, and then, the internal structure of the MEMS gyroscope is optimized to suppress the temperature drift [5].

Until now, some attempts have been made to improve the above deficiencies by studying temperature characteristics of the structure and peripheral circuit of MEMS gyroscopes [6]. For example, Nesterenko et. al. [7] presents the design and simulation of a microelectromechanical gyroscope that simultaneously determines two components of angular velocity. In order to consider how temperature influences eigenfrequencies and informative vibrational magnitude of the micromechanical angular velocity sensor. Guo et al. [8] adopted the finite element method (FEM) to verify the feasibility of the design and to compute the performance of the system. Results show that in the range of −20 to 80 °C, the maximum relative error of the resonant frequency variation reducing from 16.3% to 3.1% indicates that this design scheme is effective in overcoming the temperature effect of this kind of gyroscope. Fu et al. [9] designed a constant trans-conductance high-linearity amplifier that could realize the low-phase drift and low-amplitude drift interface circuit at all temperatures, and the test result shows that the zero-point drift is lower than 30°/h (1-sigma) at the temperature range from −40 to 60 °C after three-order compensation made by the driving force. Cao et al. [10] presents the bandwidth-expanding method with a wide temperature range for sense mode coupling with a dual-mass MEMS gyroscope, and the turntable test results show that the sensing closed loop works stably in a wide temperature range (from 40 to 60 °C), and the bandwidth values are 107 and 97 Hz. Cao et al. [11] also proposed a sense mode closed-loop method for a dual-mass MEMS gyroscope based on the bipole temperature compensation method, and the bias temperature coefficient decreased from 9.534×10^−4°C to 5.991×10^−5°C. Cao et al. [12] demonstrated a closed-loop controlling system for the MEMS gyroscope sense mode, which can reduce bias temperature coefficients from 10.59° h/°C to 3.59° h/°C. Wang et al. [13] developed a digital output disk resonator gyroscope (DRG) on-chip temperature compensation method by using the virtual temperature sensor to complete the on-chip temperature compensation of the DRG angular velocity output, and the second-order compensation realizes the scale factor change of 40 ppm/°C and zero-output change of 27°/h over the full temperature range varying from −40 to 60 °C.

The above research proved that the accuracy of the MEMS gyroscope can be improved by improving the peripheral hardware circuit of the MEMS gyroscope’s structure; however, this type of method has the disadvantages of a long research period, cumbersome research content, and unstable results, which have an adverse effect on the accuracy of the MEMS gyroscope. Therefore, many scholars have devoted to improving the above deficiencies by studying temperature modal of MEMS gyroscopes. Shen et al. [14] proposed a novel multiple inputs/single output model based on the genetic algorithm (GA) and Elman neural network (Elman NN), and the comparison results between the traditional temperature-based model and the proposed multi-input/single output model show that the Allan variance coefficients are decreased, specifically, quantization noise (Q) from 0.0012 to 8.37×10^−4, N from 1.19×10^−3 to 5.01×10^−6, bias instability (B) from 2.69×10^−4 to 1.24×10^−4, K from 0.0013 to 5.91×10^−4, R from 9.48×10^−4 to 3.27×10^−4; the comparison results between the GA–Elman NN and traditional Elman NN show that the modeling accuracy is effectively improved, specifically in Allan variance coefficients, Q from 0.0051 to 8.37×10^−4, random walk of angular velocity (N) from 3.08×10^−5 to 5.01×10^−6, B from 8.02×10^−4 to 1.24×10^−4. Wang et al. [15] introduced a new method
that includes the radial basis function neural network (RBF NN), RBF NN based on the genetic algorithm (GA), and RBF NN based on GA with Kalman filter (KF). The experimental results proved the correctness of these three methods, and the MEMS gyroscope temperature energy influence drift is compensated effectively. Shen et al. [16] also developed a temperature error processing method for a dual-mass MEMS gyroscope based on a multi-scale parallel model. Compared with the conventional serial model, the proposed parallel model can eliminate the temperature errors more effectively with each parameter of Allan analysis improved. Specifically, the factor Q reduced from 0.035 to $9.93 \times 10^{-4}$, the factor N reduced from $2.13 \times 10^{-3}$ to $7.94 \times 10^{-6}$, and B reduced from $5.28 \times 10^{-4}$ to $4.79 \times 10^{-4}$. Wu et al. [17] created an adaptive multi-scale method based on the combination of a generalized morphological filter (CGMF) for denoising of the output signal from a MEMS gyroscope. The proposed algorithm has a better performance in ARW and bias instability, which shows better advantages in gyroscope denoising. Shi et al. [18] described a double hidden layer long short-term memory (LSTM), which is presented to predict temperature data for the gyroscope (including single point and period prediction), and the LSTM network can be used to predict the temperature (time series data). Ma et al. [19] proposed a parallel denoising model based on PE-ITD and SA-ELM, and the result shows that the bias stability improves from $0.1874^\circ/h$ to $1.599 \times 10^{-3}^\circ/h$ at a temperature varying from $-40$ to $60 \, ^\circ C$ (enhanced by 99.1%), which indicated that the designed new method is more accurate and effective. In Chang’s [20] study, a denoising and temperature drift compensation parallel model method based on wavelet transform and forward linear prediction (WFLP) and support vector regression based on the cuckoo search algorithm (CS-SVR) was proposed, for the sake of decreasing the effects of noise and temperature error on the measurement accuracy of MEMS gyroscopes. Experimental results demonstrated that the proposed method can decrease noise and compensate for the temperature error effectively. Gu et al. [21] proposed a bias drift self-calibration method for MEMS gyroscopes based on noise-suppressed mode reversal without the modeling of a bias drift signal. The experimental results show that the proposed method is feasible and could achieve a better performance than the typical mode reversal. Song et al. [22] established a real-time wavelet denoising method used for the error compensation of MEMS gyroscopes, and the results show that the $1\sigma$ standard deviation of the residual signal is $0.0243^\circ/s$, and the $1\sigma$ standard deviation of the residual signal is $0.0175^\circ/s$ after the noise reduction by the proposed method. Ding et al. [23] proposed an improved variational mode decomposition-wavelet threshold denoising (WTD) method to enhance the performance of MEMS gyroscopes. For the static signal, the mean square error (MSE) of the proposed method reduces by 10.1%, and the signal-to-noise ratio (SNR) increases by 14.2%. For the dynamic signal, the MSE of the proposed method decreases by 16.9%, and the SNR increases by 18.8%. Zhang et al. [24] proposed a serial-parallel estimation model (SPEM)-based sliding mode control (SMC) of MEMS gyroscopes. The simulation results show that the proposed controller obtains higher tracking accuracy and faster convergence, while the compound nonlinearity approximation has higher precision, and the proposed scheme is verified by simulations. Wang et al. [25] proposed a new model based on fusing an unscented Kalman filter (UKF) with support vector regression (SVR) optimized by the adaptive beetle antennae search (ABAS) algorithm. The experimental results show that the noise intensity (NI) and Durbin–Watson (DW) value of the proposed scheme in terms of the compensation accuracy for random drift data reduces and improves by 28.57% and 9.06%, respectively, compared with the conventional method. Huang et al. [26] proposed a calibration method based on deep learning concentrating on MEMS IMU gyroscopes. In this method, the output model of a MEMS IMU gyroscope is constructed by using the temporal convolutional network. The experimental results show that the attitude and position accuracy obtained by the inertial navigation solution using the compensated gyroscope data are improved compared with the existing MEMS sensor error compensation method based on deep learning, which proves that the proposed method can effectively and accurately calibrate the gyroscope error. Although there is much research on the output error of MEMS gyroscopes, tem-
temperature compensation models, and corresponding filtering algorithms, these algorithms still have many shortcomings in terms of computing speed and filtering capabilities. The temperature compensation proposed in this study is based on compensation temperature models and filtering algorithms to make up for this defect and improve the accuracy of the MEMS gyroscope.

In this study, a LSM6DS3 MEMS gyroscope with a novel structure and model was created. An improved VMD method based on convolutional neural networks—long short-term memory (CNN-LSTM) and particle swarm optimization—support vector machines (PSO-SVM) was proposed to deal with the temperature error. The output data of the X axis under the ranges of temperature from –40 to 60 °C were discussed. Furthermore, the proposed methods were compared with the Allan variance analysis method corresponding to the performance of the MEMS gyroscope to improve practicability and significance of the suggested methods.

2. Structure of MEMS Gyroscope

As is shown in Figure 1, the structure of this MEMS gyroscope is composed of an anchor, Coriolis mass, drive frame, sense frame, drive comb, sense comb sense spring, and drive spring. The structure of this MEMS gyroscope can be divided into two modes, namely drive mode and sense mode, which can be described as a spring–mass–damping second order system. The drive mode and the sense mode have one DOF along the X direction and Y direction, respectively. Each axis has two sections, which are stiffness and damping. As is described in Figure 1, Coriolis mass has two DOF on both the X axis and Y axis. In general, MEMS gyroscopes work on the drive mode by driving stiffness and drive damping, and the sense stiffness and sense damping are driven by the sensing mode to detect angular velocity when the Coriolis mass is subjected to an angular velocity of Z axis. The dynamic system can be represented by the simplified equations [27,28].

\[
m\ddot{x} + d_x \dot{x} + k_x x = F_d \sin(\omega_d t) \tag{1}
\]

\[
m\ddot{y} + d_y \dot{y} + k_y y = -2m\Omega_z \dot{x} \tag{2}
\]

where \(m\) is the main mass of this MEMS gyroscope, which is used by the drive mode and sense mode. \(F_d\) is drive force amplitude, \(\omega_d\) is drive force angular frequency, and \(\Omega_z\) is angular rate around the Z axis.

![Figure 1. The structure of the MEMS gyroscope.](image-url)
By solving Equation (1), the displacement of the MEMS gyroscope in the drive direction can be described as:

\[
x(t) = \frac{2F_{d}L_{d}}{(m_{d}^{2} - m_{w}^{2})^{2} + 4L_{d}^{2}m_{d}^{2}w_{d}^{2}}e^{-\zeta_{d}w_{d}t} \cos(w_{d}t\sqrt{1 - \xi_{d}^{2}}) \\
+ \frac{F_{d}L_{d}w_{d}^{2} + \omega_{d}^{2} - \omega_{w}^{2}}{m_{w}}e^{-\zeta_{w}w_{w}t} \cos(w_{w}t\sqrt{1 - \xi_{w}^{2}}) \\
+ \frac{F_{d}/m_{w}}{w_{d}^{2} \sqrt{1 - \xi_{d}^{2}}(m_{d}^{2} - m_{w}^{2})^{2} + 4L_{d}^{2}m_{d}^{2}w_{d}^{2}}e^{-\zeta_{w}w_{w}t} \sin(w_{w}t - \varphi_{1})
\]

By solving Equation (2), the displacement of the MEMS gyroscope in the sense direction can be described as:

\[
y(t) = \frac{2m_{w}L_{d}A_{1}w_{d}(2\zeta_{d}^{2}w_{d}^{2}w_{w}^{2}t + \varphi_{1} + (w_{d}^{2} - w_{w}^{2}) \cos\varphi_{1})/m_{w}}{(w_{d}^{2} - w_{w}^{2})^{2} + 4L_{d}^{2}m_{d}^{2}w_{d}^{2}}e^{-\zeta_{w}w_{w}t} \cos(w_{w}t\sqrt{1 - \xi_{w}^{2}}) \\
+ \frac{2m_{w}L_{d}A_{1}w_{d}(w_{d}^{2} - 3w_{w}^{2}) \cos\varphi_{1} + w_{d}^{2}(2\zeta_{d}^{2}w_{d}^{2}w_{w}^{2} + w_{d}^{2} - w_{w}^{2}) \sin\varphi_{1})/m_{w}}{w_{d}^{2} \sqrt{1 - \xi_{d}^{2}}(m_{d}^{2} - m_{w}^{2})^{2} + 4L_{d}^{2}m_{d}^{2}w_{d}^{2}}e^{-\zeta_{w}w_{w}t} \sin(w_{w}t\sqrt{1 - \xi_{w}^{2}}) \\
+ \frac{2m_{w}L_{d}A_{1}w_{d}/m_{w}}{w_{d}^{2} \sqrt{1 - \xi_{d}^{2}}(m_{d}^{2} - m_{w}^{2})^{2} + 4L_{d}^{2}m_{d}^{2}w_{d}^{2}}e^{-\zeta_{w}w_{w}t} \cos(w_{w}t - \varphi_{1} - \varphi_{2})
\]

Figure 2 illustrates that the circuit structure consists of the driving mode and sensing mode. The driving mode is composed of the MEMS gyroscope structure, amplifying phase, rectifier, low pass filter (LPF), DC voltage component, multiplier, and adder. The sense mode includes four different kinds of low pass filter (LPF), two different signal amplifiers, and two multipliers, which represent resonant frequency of driving mode and quality factor of driving mode, respectively. As for the drive mode, all driving mode loops form a self-excited oscillation loop so that the driving mode of the MEMS gyroscope always works in the resonance mode. As for the sense mode, this MEMS gyroscope adopts an open-loop sense mode; thus, the influence of the quadrature signals on circuit accuracy cannot be eliminated.

![Figure 2. Circuit structure of MEMS gyroscope.](image-url)

Figure 2 also displays the curve of the voltage of the MEMS gyroscope drive signal, sense signal and the quadrature signal along with time. The drive signal, sense signal, and quadrature signal are all varying by sine or cosine along with time. Due to the existence of the quadrature signal, the output error of the gyroscope is large, and the overall accuracy of the gyroscope is low. To address this problem, some software methods were proposed in this study to improve the accuracy of the MEMS gyroscope.
Figures 3 and 4 show that the MEMS gyroscope has two corrected modes and four irrelevant modes. One of the corrected modes is the driving mode (Figure 3a) with a simulated resonance frequency of 11,034.6 Hz, and the other is the sensitive mode (Figure 3b) with a simulated resonance frequency of 11,035.2 Hz. In the driving mode, the Coriolis mass moves in resonance along the X-axis. In the sensitive mode, the same Coriolis mass moves in resonance along the Y axis because this structure is a fully decoupled structure, and the X-axis driving mode and the Y-axis sensitive mode do not interfere with each other. Figure 4 shows that the frequency difference between the driving mode and the sensitive mode is 0.6 Hz, and each mode is a fully decoupled mode independent of each other.

![Figures 3 and 4](image1)

**Figure 3.** The vibration mode of MEMS gyroscope: (a) drive mode; (b) sense mode; (c) irrelevant mode; (d) irrelevant mode; (e) irrelevant mode; (f) irrelevant mode.

![Figures 3 and 4](image2)

**Figure 4.** Frequency difference between driving mode and sensitive mode.

3. Algorithms

3.1. The Algorithm of VMD

The variational modal decomposition (VMD) algorithm was proposed by K. Dragomiretsky, who argued that a complex signal is composed of sub-signals with different frequency...
bandwidths. The function of VMD is to decompose the composite signal into multiple sub-modes according to the frequency bandwidth. The VMD method completes the decomposition of the signal and the acquisition of signal components under the variational framework and decomposes the original signal adaptation by constructing and solving the variational constraint problem. In the application of fault signal processing, it can effectively decompose modal components with different center frequencies and bandwidths according to the frequency domain characteristics of the signal; meanwhile, it is not easily affected by frequency changes and has good noise robustness.

The first step is to construct the variational problem:

VMD is used to process fault signals and is mainly responsible for decomposing the signal into modal components with their respective optimal center frequencies and limited bandwidths according to the actual situation and given values. The objective function is the minimum sum of the bandwidths of the decomposed modal components.

The constrained variational problem model is as follows:

\[
\min \left\{ \{u_k\}, \{w_k\}, \lambda \right\} \left\{ \sum_k \left[ \delta(t) + \frac{j}{\pi t} u_k(t) e^{-j\omega_k t^2} \right] \right\}
\]

where the K modal components are expressed as \(\{u_k\} = \{u_1, \ldots, u_K\}\), the center frequency is expressed as \(\{w_k\} = \{w_1, \ldots, w_K\}\), and \(\delta(t)\) is the unit impulse function.

The second step is to solve the variational problem. After completing construction of the first step, the next step is to solve the above equation and to convert it into an unconstrained problem. Two parameters, quadratic penalty factor \(\alpha\) and Lagrange multiplication operator \(\lambda(t)\), need to be introduced into the solution. The robustness of quadratic penalty factor \(\alpha\) noise is strong, and it can ensure that the fault signal can be reconstructed well even if it is disturbed by the surrounding environment and other factors. The introduction of the Lagrange multiplication operator \(\lambda(t)\) can significantly change the constraints of the conditions of the variational problem and turn it into an unconstrained problem. The extended Lagrange function expression obtained by the combination of the two is:

\[
L(\{u_k\}, \{w_k\}, \lambda) = \alpha \sum_k \left[ \delta(t) + \frac{j}{\pi t} u_k(t) \right] e^{-j\omega_k t^2} + f(t) - \sum_k u_k(t) + \frac{\lambda(t)}{2} \]

The third step is to find the saddle point of the extended Lagrange function above, that is, the optimal solution of the expression of Equation (6). Using the alternating direction multiplier algorithm to solve the original minimization problem, it is necessary to continuously update \(u_k^{n+1}, w_k^{n+1}\) and \(\lambda^{n+1}\). The expression of \(u_k^{n+1}\) is shown below:

\[
u_k^{n+1} = \arg \min_{u_k \in X} \left\{ \alpha \left\| \delta(t) + \frac{j}{\pi t} u_k(t) \right\|^2_2 + \left\| f(t) - \sum_i w_i(t) + \frac{\lambda(t)}{2} \right\|^2_2 \right\}
\]

For simplicity, let \(w_k = w_k^{n+1}\), then \(\sum_t u_i(t) = \sum_{t \neq k} u_i(t)^{n+1}\), and Equation (8) is processed into the frequency domain space form by transformation as follows:

\[
\hat{u}_k^{n+1} = \arg \min_{\hat{u}_k, u_k \in X} \left\{ \alpha \left\| \hat{u}_k (1 + \sgn(w + w_k)) u_k (w + w_k) \right\|^2_2 + \left\| f(w) - \sum_i \hat{u}_i (w) + \frac{\lambda (w)}{2} \right\|^2_2 \right\}
\]
where \( \hat{\lambda}(5) \) is the Wiener filter of the current residual \( \hat{u} \), and \( \hat{\lambda}(n) \) is the Wiener filter of the current residual \( \hat{u} \).

Using the conjugate symmetry of the signal, the non-negative frequency interval of Equation (9) is integrated as follows:

\[
\hat{u}^n_{k+1} = \arg\min_{\hat{u}_k, \hat{u}_k} \left\{ \int_0^\infty 4\alpha(w - w_k)^2 \hat{u}_k(w)^2 + 2 \left| \hat{f}(w) - \sum_i \hat{u}_i(w) + \hat{\lambda}(w) \right| dw \right\}
\]  

(10)

The equation is quadratically optimized in the positive frequency range to obtain

\[
\hat{u}^n_{k+1} = \frac{\hat{f}(w) - \sum_{i \neq k} \hat{u}_i(w) + \hat{\lambda}(w)}{1 + 2\alpha(w - w_k)^2}
\]  

(11)

where \( \hat{u}^n_{k+1}(w) \), \( \hat{f}(w) \), and \( \hat{\lambda}(w) \) are the Fourier transforms corresponding to \( u_k^{n+1}, f(w) \) and \( \lambda(w) \), respectively. \( \hat{u}^n_{k+1}(w) \) is the Wiener filter of the current residual \( \hat{f}(w) - \sum_{i \neq k} \hat{u}_i(w) + \hat{\lambda}(w) \).

In the same way, the center frequency \( \hat{w}^n_{k+1} \) can be expressed as:

\[
\hat{w}^n_{k+1} = \arg\min_{\hat{w}_k} \left\{ \partial_t \left[ \left( \delta(t) + \frac{i}{\pi t} \right) \ast u_k(t) \right] e^{-i\hat{w}_k t^2} \right\}
\]  

(12)

Equation (13) is obtained by converting to the frequency domain interval.

\[
\hat{w}^n_{k+1} = \arg\min_{\hat{w}_k} \left\{ \int_0^\infty (w - w_k)^2 \hat{u}_k(w)^2 dw \right\}
\]  

(13)

The updated method of solving the center frequency is expressed in Equation (14).

\[
\hat{w}^n_{k+1} = \frac{\int_0^\infty w\hat{u}_k(w)^2 dw}{\int_0^\infty \hat{u}_k(w)^2 dw}
\]  

(14)

The VMD algorithm decomposes the original signal through loop iteration and obtains the corresponding number of effective IMF components by presetting the value of K in the process of decomposition. On variational problem solving, the model is first transformed from the time domain to the frequency domain, the parameters \( u_k \) and \( \hat{w}_k \) are continuously optimized and updated, and finally, the inverse Fourier transform is carried out to the time domain to obtain the center frequency of each mode.

The complete VMD algorithm follows these six steps with the detailed process illustrated in Figure 5.

1. Initialize \( \{ \hat{u}_k \}, \{ \hat{w}_k \}, \hat{\lambda} \), let \( n = 0 \).
2. \( n = n + 1 \), start the cycle.
3. Update \( u_k \) according to Equations (3)–(17), update \( \hat{w}_k \) according to Equations (3)–(20).
4. \( k = k + 1 \), repeat step (3). If \( k = K \), the cycle ends, and if \( k \neq K \), the cycle continues.
5. \( \lambda \) is updated according to the following expression.

\[
\hat{\lambda}^{n+1}(w) \leftarrow \hat{\lambda}^n(w) + \tau(\hat{f}(w) - \sum_k \hat{u}^{n+1}_k(w))
\]  

(15)
Steps (2) to (5) are repeated, with the given precision convergence criterion, and 
\[ \sum_k \frac{\| \hat{u}_{n+1}^k - \hat{u}_n^k \|}{\| \hat{u}_n^k \|} < \varepsilon \]
is judged whether it is satisfied. If so, turn to this step and stop the iteration to obtain IMF components. Otherwise, skip to step (2).

Figure 5. VMD algorithm flow chart.

3.2. The Algorithm of CNN-LSTM

The advantages of both CNN and LSTM networks are combined to propose a novel algorithm. First of all, the CNN network performs well in extracting N-gram features at different positions in neurons, which can be extracted by convolution operations, so that it can be used to identify noise in gyro output signals. The LSTM network can deal with the noise of any length and extract the dependencies between the noise. After combining the two, it can benefit from the advantages of the two networks and identify different noise in the gyro output signal more accurately and quickly. The designed CNN-LSTM algorithm takes the gyro output signal as input, which is connected to a convolution module after passing through the embedding layer; then, the feature vector extracted by the CNN network is reduced in dimension through a maximum pooling layer and is sent to an LSTM module to extract features. Finally, the Sigmoid activation function is used to classify the noise of the gyro output signal to determine whether it is a useful signal. Below, the specific design of each layer in the network is described in detail. The overall structure of the network is shown in Figure 6, and the specific algorithm flow is shown in Figure 6.
Figure 6. CNN-LSTM algorithm flow chart.

When the output of the MEMS gyroscope completes the data preprocessing, the CNN module is first entered. Since the gyro signal vector we process is a one-dimensional vector, one-dimensional convolution is chosen for processing. This step is mainly meant to extract the relevant features in the output of the gyroscope through convolution operation. $x \in \mathbb{R}^d$ is set to represent the one-dimensional vector with the drift of $d$ corresponding to the i-th time point in the output of the gyroscope, and $x \in \mathbb{R}^{l+d}$ represents the input gyro signal, where $l$ represents the length of the output of the gyroscope. Then, the window vector $w_j$ of each position $j$ in the signal corresponding to the continuous k-length of the output of the gyroscope can be expressed as:

$$w_j = [x_{j}, x_{j+1}, \ldots, x_{j+k-1}]$$  \hspace{1cm} (16)

Then, through the filter in the convolutional layer, the feature vector corresponding to the window vector can be calculated, and the calculation method is shown in Equation (17).

$$c_j = f(w_j \odot p + b)$$  \hspace{1cm} (17)

where $\odot$ represents the element-wise multiplication of feature vectors, $b \in \mathbb{R}$ is a bias term, and $f$ is a nonlinear mapping function. In the research process, the function is set as the ReLU activation function, which is defined as Equation (18).

$$f(x) = \max(0, x)$$  \hspace{1cm} (18)

During the calculation, if the element in the vector is a regular element, it returns the element; otherwise, it returns to 0. The specific browsing learning process is shown in Algorithm 1.

\begin{algorithm}
\caption{CNN model}
\begin{algorithmic}
\State \textbf{Input}: $x \in \mathbb{R}^{l+d}$
\State \textbf{Output}: $\hat{c}$
\For {each position $j$ in the signal, perform}
\State $w_j = [x_{j}, x_{j+1}, \ldots, x_{j+k-1}]$
\State $c_j = \text{ReLU}(w_j \odot p + b)$
\EndFor
\State $c = [c_1, c_2, \ldots, c_{l+k-1}]$
\State $\hat{c} = \max(c)$  //Max pooling operation
\end{algorithmic}
\end{algorithm}

The feature vector obtained by the convolution layer is the high-dimensional feature of the output signal of the gyro. In order to filter out the redundant noise, only the important features in the output signal of the gyro are retained to avoid the trained network caused by the noise information in the text. In the study, a max pooling layer is added after the CNN module to decrease the dimension of the feature vector, which can also reduce the computational cost of the network. The CNN network has a very good performance in extracting relevant features from the gyro output signal data. However, it cannot correlate current content with past content in conjunction with contextual information in the gyro output signal. Therefore, another deep learning network LSTM is added to the structure to complete the learning of associated features.

The basic structure of a LSTM network consists of a series of repeating units at each time step. In each unit, at time step $t$, an information storage part $c_t$ and three gate functions, namely input gate $i_t$, output gate $o_t$ and forget gate $f_t$, are used to regulate and manage the information flow of each unit in the LSTM network and to decide how to update the information stored in the current storage unit $c_t$ and the current hidden state $h_t$ of the unit. The relevant calculation function of each unit in the LSTM network module is listed in Equations (19)–(24).

$$i_t = \sigma(W_i[h_{t-1} + b_i])$$
$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$
$$q_t = \tanh(W_q[h_{t-1}, x_t] + b_q)$$
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot q_t$$
$$h_t = o_t \odot \tanh(c_t)$$

where $x_t$ is the input feature in the LSTM network unit, which is extracted by the above CNN module. $\sigma$ represents the Sigmoid function, $\odot$ represents the element-by-element multiplication of feature vectors, $W$ and $b$ represent the weight matrix and offset vector during training, respectively. In the structure of the proposed model, an LSTM module is placed directly after the max pooling layer, containing 64 LSTM units, which use a dropout layer of 0.2 as a regularization parameter to prevent the model from overfitting. The specific training and learning processes are shown in Algorithm 2.

**Algorithm 2. LSTM mode**

**Input:** Signal with noise  
**Output:** Signal  
**For each time step $t$, perform**

1. $i_t = \sigma(W_i[h_{t-1} + b_i])$
2. $f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$
3. $q_t = \tanh(W_q[h_{t-1}, x_t] + b_q)$
4. $o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$
5. $c_t = f_t \odot c_{t-1} + i_t \odot q_t$
6. $h_t = o_t \odot \tanh(c_t)$

End

The dense layer is the last layer of the entire model. That is, the fully connected layer in the neural network is used to output the result and to classify the gyro output signal according to the output of the LSTM layer. Since the dataset used for training in this paper divides the gyro output signal into two categories, our classification model is binary. A fully connected layer and sigmoid function are used to provide 0 or 1 predictions for
two classes (useful signal and noise), where the sigmoid function is a logical function whose return value is between 0 and 1, as defined in Equation (25).

\[ f(x) = \frac{1}{1 + e^{-x}} \]

(25)

3.3. The Algorithm of PSO-SVM

The PSO algorithm, proposed by Kennedy and Eberhardt in 1995, is an intelligent biological algorithm. The algorithm takes advantage of mineral populations in nature and uses their behavior to first solve complex optimization problems with the help of environmental services. Compared with the genetic algorithm, the PSO algorithm has the advantages of speed dissipation, easy implementation, and high precision.

PSO also introduces the thorny problem of how to perform a random search in a D dimensional space with the goal of solving the problem of maximizing or minimizing the objective function, where \( X = (X_1, X_2, \ldots, X_n) \), \( x_i = (x_{i1}, x_{i2}, \ldots, x_{iD}) \), and \( v_i = (v_{i1}, v_{i2}, \ldots, v_{iD}) \) represent a population composed of \( n \) particles, the position vector of a single particle \( i \), and the velocity vector, respectively. When the particle \( i \) searches the D dimensional space, the local optimal solution is the optimal position searched for \( p_{best_i} = (p_{best_{i1}}, p_{best_{i2}}, \ldots, p_{best_{iD}})^T \), and the global optimal solution is the optimal position \( g_{best} = (g_{best_1}, g_{best_2}, \ldots, g_{best_D})^T \) searched by the entire particle swarm.

In the iterative process, the velocity of each particle is modified and determined according to the position of the local optimal particle, the position of the global optimal particle, and the velocity and position of the particle itself. The local optimal position is the optimal position reached by each particle in the iterative optimization process. The particle velocity calculation and position calculation results are shown in Equations (26) and (27).

\[
v_i(t + 1) = w_i(t) + c_1r_1(p_{best_i}(t) - x_i(t)) + c_2r_2(g_{best}(t) - x_i(t)) \tag{26}
\]

\[
x_i(t + 1) = x_i(t) + v_i(t + 1) \tag{27}
\]

where \( i = 1, 2, \ldots, N \), \( N \) is the average number, \( w \) is the inertia coefficient, and its value is not negative. When its value is relatively large, the global search ability is strong while the local search ability is weak. When its value is small, the global search is weak, the local search is strong, and the best search results can be obtained. \( it_{max} \) is the maximum number of iterations, \( w_{ini} \) is the initial inertia weight, the typical weight is \( w_{ini} = 0.9 \). \( w_{end} \) is inertia weight when the iteration \( w_{end} \) reaches the maximum algebra, and the typical weight is \( w_{end} = 0.4 \). The local and global learning factors are \( c_1, c_2 \) with the range of \( 0 \leq c_1, c_2 \leq 2 \), (usually taken as \( c_1 = c_2 = 2 \)). \( r_1, r_2 \) are two random numbers in the range \((0, 1)\), where the particle’s position and velocity are limited to \([-x_{max}, x_{max}], [-v_{max}, v_{max}]\), \( p_{best} \) is the local optimal position of the particle, and \( g_{best} \) is the full optimal position. The basic PSO algorithm steps are summarized below:

Step1: For each particle in the population, its fitness needs to be calculated.
Step2: For each particle, the best fitness value passed by it is compared with the fitness value. If it is better, it is treated as a locally optimal particle.
Step3: For each generation of optimal particles, the global optimal particle is compared with its fitness value, and if it is better, it can be taken as the global optimal particle.
Step4: The speed and position of the particles are adjusted according to the above formula.
Step5: If the corresponding conditions are not met, go back to step 1.

The termination condition of the algorithm iteration is that the optimal position searched by the particle swarm reaches the minimum fitness threshold or the algorithm has iterated to the set maximum number of iterations.

The SVM algorithm is a new machine learning method based on the VC theory of statistical learning and the principle of structural risk minimization. Its core idea is to transform and map the nonlinear data into a high-dimensional linear space to meet the maximum
classification distance, so that the classification line can correctly separate the two types of samples, and then, the required optimal classification hyperplane is obtained. As shown in Figure 7, assuming that the given dataset is \( \{x_i, y_i\}, i = 1, 2, \ldots, N, y_i \in \{-1, +1\}, x_i \in \mathbb{R}^d, \) two triangles and five pointed stars are used to represent two samples on the plane.

![Schematic diagram of support vector machine.](image)

**Figure 7.** Schematic diagram of support vector machine.

The SVM introduces the slack variable \( \xi_i \geq 0 \) and penalty coefficient \( C(C > 0) \). Then, the solution formula of the hyperplane is:

\[
\begin{align*}
\max & \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j (x_i, x_j) \\
\text{s.t.} & \quad 0 \leq \alpha_i \leq C, i = 1, \ldots, m \\
& \quad \sum_{i=1}^{m} \alpha_i y_i = 0
\end{align*}
\] (28)

In the formula, \( \alpha_i \) is the Lagrangian factor solved by the quadratic optimization problem to obtain the optimal classification hyperplane.

\[
K(x, x') = \exp\left( - \frac{||x - x'||^2}{\sigma^2} \right)
\] (29)

Then, the optimal classification hyperplane is:

\[
f(x) = \text{sign} \left[ \sum_{i=1}^{N} \alpha_i y_i k(x_i, x) + b \right]
\] (30)

The algorithm flow of the prediction model establishment based on PSO-SVM is shown in Figure 8. The specific steps of the model establishment are as follows:

1. **Step 1:** Obtain relevant data, determine the number of training samples and test the samples, quantify the sample data uniformly, and normalize the quantized data to the interval \([0, 1]\).

2. **Step 2:** Set the initial parameters of the PSO-SVM model, including the population size, the number of iterations, the learning factor, the inertia weight, the initial particle position and the initial particle velocity, etc.

3. **Step 3:** Evaluate the fitness of each particle and update the individual extremum and global extremum of the particle swarm.

4. **Step 4:** Judging whether the end condition is met or the terminating evolutionary algebra is reached, output the optimal penalty factor \( C \) and kernel function \( K \); if the judgment result is no, then return to step 3.

5. **Step 5:** Input the optimal penalty factor \( C \) and kernel function \( K \) into the SVM model for sample training, and obtain the globally optimal PSO-SVM model, which is used to predict the data.
3.4. The Algorithm of ELM

Given an initial training sample \((x_i, t_i)\), the input sample is denoted as \(x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^T \in \mathbb{R}^n\), and the output sample is denoted as \(t_i = [t_{i1}, t_{i2}, \ldots, t_{im}]^T \in \mathbb{R}^m\). The SLFN network model with hidden layer nodes \(L (N_0 \geq L)\) and activation function is \(g(x)\), as shown in Figure 9.

In order to make the learning objective of SLFNs minimize the output, it can be expressed in Equation (31), where \(w_i = [w_{i1}, w_{i2}, \ldots, w_{in}]^T\) represents the weight vector of the input node and the node connecting the hidden layer, \(\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^T\) represents the weight vector used to connect the output node and the hidden layer node, \(b_i\) is the bias of the \(i\)th hidden layer node, and \(w_i \cdot x_j\) represents the inner product of \(w_i\) and \(x_j\).

\[
\sum_{i=1}^{L} \beta_i g(x_i) = \sum_{i=1}^{L} \beta_i g(w_i \cdot x_j + b_i) = o_j, j = 1, 2, \ldots, N
\]  
(31)

\[
\sum_{j=1}^{N} \|o_j - t_j\|
\]  
(32)
That is, there is $\beta_i$ and $\omega_i$ to make Equation (31) true, as shown in Equation (33).

$$
\sum_{i=1}^{L} \beta_i g(w_i x_j + b_i) = t_j, j = 1, 2, \ldots, N
$$

Equation (31) can be simplified to Equations (34) and (35).

$$
H\beta = T
$$

$$
H(w_1, \ldots, w_L, b_1, \ldots, b_L, x_1, \ldots, x_L) = \begin{bmatrix}
g(w_1 x_1 + b_1) & \ldots & g(w_L x_1 + b_L) \\
\vdots & \ddots & \vdots \\
g(w_1 x_N + b_1) & \ldots & g(w_L x_N + b_L)
\end{bmatrix}_{N \times L}
$$

In Equation (35), $H$ is the output matrix of the neural network hidden layer interface, and the column of $L$ is the output matrix $H$ of the $i$ hidden layer node. For some gradient proximity algorithms, when the deployment function is infinite, the input link density and hidden layer irregularity can be randomly generated for training, which will change during training, to ensure that the output link weights $\beta$ are obtained by the least squares solution in Equation (36).

$$
\min ||H\beta - T||
$$

In the ELM neural network algorithm, random connection weights $w_i$ and hidden layer biases $b_i$ are obtained, and the output matrix $H$ of the hidden layer is determined. The ELM neural network can be trained to change into a linear system $H\beta = T$, and the output weights $\hat{\beta}$ can be determined, as shown in Equation (37).

$$
\hat{\beta} = H^+ T
$$

where $H^+$ indicates that the Moore–Penrose generalized inverse is the hidden layer output matrix $H$, and the norm $\hat{\beta}$ needs to be known as the smallest and most unique. In summary, the steps of the ELM learning method can be summarized as follows:

Step1: Given a training set $(x_i, t_i), (i = 1, 2, \ldots, N)$, the activation function is $g(x)$, the number of hidden layer nodes is $L$, the random initialization input weight is $w_i$, and the hidden layer node offset is $b_i$.

Step2: Calculate the hidden layer output matrix $H$.

Step3: Calculate the output weight matrix $\beta$. 
Due to the ELM algorithm, the output weight matrix can be obtained only by calculation, which can greatly reduce the training complexity and greatly improve the training efficiency. The core idea of the ELM algorithm is to calculate the output weight matrix of the hidden layer through the Moore–Penrose generalized inverse and to transform the training process of the traditional SLFN model into a least-squares solution problem, and the calculated solution is unique. The difference between ELM and traditional SLFNs is that ELM does not need to iteratively adjust all node parameters of the hidden layer during model building and training processes. Meanwhile, the node parameters are randomly generated during the network training process. Furthermore, when training the network model, there is no need to adjust the parameters.

4. The Improved VMD and ELM Algorithm Based on CNN-LSTM and PSO-SVM

In this paper, an improved VMD and ELM algorithm is designed and proposed based on CNN-LSTM and PSO-SVM. This optimization algorithm identifies the noise of different frequencies in the gyro output signal through VMD and then passes the high-, medium- and low-frequency noise through the CNN-LSTM. PSO-SVM algorithms are used for modeling analysis, and then, the optimized noise signal is passed through the ELM neural network to establish a temperature compensation model, and finally, the optimized gyro output signal is obtained. Figure 10 describes the specific process of this improved method:

1. Preprocess the original signal output by the MEMS gyroscope, and then, perform frequency division processing on the preprocessed signal through the VMD algorithm to obtain three groups of signals, which are high-frequency noise signal \( \eta_{\text{high}}(t) \), intermediate-frequency temperature noise signal \( \eta_{\text{medium}}(t) \) and low-frequency noise signal \( \eta_{\text{low}}(t) \).

2. The high-frequency noise signal \( \eta_{\text{high}}(t) \) is discarded, and then, the low-frequency noise signal \( \eta_{\text{low}}(t) \) is modeled through the CNN-LSTM algorithm to establish the high-frequency signal compensation model and to obtain the modeled signal \( \eta_{\text{low}}^2(t) \).

3. The intermediate-frequency temperature noise compensation model is established by PSO-SVM, the intermediate-frequency temperature noise signal \( \eta_{\text{medium}}^2(t) \) is compensated by the PSO-SVM algorithm, and the compensated signal \( \eta_{\text{medium}}^2(t) \) is obtained.

4. The temperature \( T \), temperature change rate \( dT \), low-frequency noise compensation signal \( \eta_{\text{low}}^2(t) \), and intermediate-frequency temperature noise compensation signal \( \eta_{\text{medium}}^2(t) \) are used as the input layer of the ELM neural network, and a temperature compensation analysis model is established. Finally, the final temperature compensation output of the gyroscope is obtained at the output layer of the ELM neural network signal.

Figure 10. The specific process of the improved method.

5. Experiment and Analysis

5.1. Experiment Process

The MEMS is kept on the temperature-controlled oven in order to output the signal of the MEMS vibration gyroscope, which is not influenced by outside temperature variation. Then, the ranges of temperature and temperature rate are set from \(-40\) to \(60\) °C and \(1\) °C/min, respectively. Firstly, the initial temperature should be set as \(60\) °C and kept for an hour in order to ensure that the inside temperature of the MEMS vibration gyroscope
is stable at 60 °C. Secondly, the temperature is reduced at the speed of −1 °C/min into −40 °C, and the temperature is kept for an hour to make sure the inside temperature of the MEMS vibration gyroscope is −40 °C. The temperature-controlled oven should stay at 10 °C for an hour in order to collect the output of the MEMS vibration gyroscope and to ensure that the inside temperature of the MEMS vibration gyroscope is stable and equal to the oven temperature. The equipment of the temperature test and the process of the temperature experiment are shown in Figures 11 and 12 [29–31].

**Figure 11.** Equipment of the temperature test.

**Figure 12.** Process of the full relative humidity experiment.

Based on the designed full-temperature experiment, data were collected at fixed points, the test was repeated for 10 days, ten sets of data were collected, and the sixth set of data with the average value of Allan variation through 10 sets of data was used for analysis. The output data from −40 to 60 °C and from 60 to −40 °C are shown in Figures 13 and 14, respectively.
Figure 13. The output of the gyroscope based on the X axis from −40 to 60 °C.

Figure 14. The output of the gyroscope based on the X axis from 60 to −40 °C.

5.2. Data Analysis

It can be seen from the decomposition diagram that the output signal consists of a low-frequency noise signal, intermediate-frequency noise signal, and high-frequency noise signal. The low-frequency noise signal, intermediate-frequency noise signal, and high-frequency noise are extracted through VMD decomposition, and eight natural mode functions are obtained (IMF1–IMF8), which represent different characteristics of different noise signals in the output signal. If each IMF is processed, the calculation consumes a lot, and it is easy to destroy the static information, resulting in errors; thus, this entropy is used for the modal function. Figure 15a shows the IMF1–IMF8 based on VMD from −40 to 60 °C, and Figure 15b shows the IMF1–IMF8 based on VMD from 60 to −40 °C. According to the sequence autocorrelation and complexity, the IMF is divided into high-frequency noise signal $\eta_{\text{high}}(t)$, intermediate-frequency temperature noise signal $\eta_{\text{medium}}(t)$, and low-frequency noise signal $\eta_{\text{low}}(t)$. 
If the sample entropy value of the IMF is greater than 0.6, it is a pure noise term that does not contain any useful signals, and it can be deleted directly. If the sample entropy value of the IMF is between 0.3 and 0.5, it is intermediate-frequency temperature noise, which also includes effective temperature characteristics and noise signals. The PSO-SVM method is used to deal with this part of the noise, because it is the most important part of the temperature compensation model. If the signal sample entropy value is between 0.2 and 0.3, the signal is a low-frequency noise signal, using CNN-LSTM to build the compensation model. The specific classification of high-frequency noise signal $\eta_{\text{high}}(t)$, intermediate-frequency temperature noise signal $\eta_{\text{medium}}(t)$ and low-frequency noise signal $\eta_{\text{low}}(t)$ is shown in Figure 16.

Figure 15. The IMF1-IMF8 based on VMD of (a) −40 to 60 °C and (b) 60 to −40 °C.
The IMF2 and IMF3 sequences decomposed by VMD are selected as the training and test datasets of CNN-LSTM, and the operating parameters of CNN-LSTM are set. The specific parameters are shown in Table 1. After repeated experiments and network optimization, the RMSE error value and prediction data are shown in Figures 17 and 18. Among them, Figure 17 shows the low-frequency noise filtering and optimization during the heating process (−40–60 °C), and Figure 18 shows the low-frequency noise filtering and optimization during the cooling process (60–−40 °C).

**Table 1.** CNN-LSTM specific operating parameters.

| Parameter                  | Specific Operation |
|----------------------------|--------------------|
| Number of CNN-LSTM modules | 4                  |
| Input                      | Noise              |
| Data normalization         | Min–Max            |
| Number of LSTM units       | 5                  |
| Number of LSTM layers      | 2                  |
| Softmax layer              | Softmax            |
| Loss function              | MSE                |
| Number of iteration rounds | 500                |

**Figure 16.** Noise classification based on SE-VMD and CNN-LSTM.

**Figure 17.** Noise denoising based on CNN-LSTM (−40 °C–60 °C).
SVM is used to train temperature noise data and to improve the relatively better $C$ and $\gamma$ of the sexual particle group optimization algorithm of SVM. The range of $C$ and $\gamma$ are set from 0 to 10, the value of the inertial factor $W$ is 0.5, the values of the learning factors $C_1$ and $C_2$ are set to 1.46, the total number of particles is set to 100, and the number of iterations is set to 50. The initial state and the end of the iteration of the searches of particle group optimization algorithms are shown in Figure 19. It can be seen from the figure that in the initial state, PSO randomly generates several particles. Each particle represents a set of $C$ and $\gamma$ values, and the relatively better parameters are finalized by continuous iteration updates. Ten tests are conducted under the setting conditions, and then, the average value is taken as the final result. The relatively better parameter value $C$ under the setting conditions is 7.20, and $\gamma$ is 6.75. The support vector machine model in this parameter is relatively better. Figure 20 shows the temperature noise compensation based on PSO-SVM. Among them, the compensation can remove the effects of temperature noise on the output signal of the PSO-SVM with good compensation effects.
Low-frequency noise after CNN-LSTM treatment, temperature noise after PSO-SVM treatment, and normal value noise are reconstructed through the ELM algorithm. The input signal is the low-frequency noise after CNN-LSTM treatment, the temperature noise after the PSO-SVM treatment, constant noise, temperature, and temperature change rate. The reconstructed output signal is shown in Figures 21 and 22. It can be concluded that the method used in this study can dramatically reduce the impact of temperature on the performance of the gyroscope and that it greatly improves the output accuracy of the gyroscope.
Finally, the prominent feature of Allan variance is that it can be easy to represent and identify various sources of error and the contribution of the whole noise statistical characteristics as it has the advantages of easy calculation and easy separation. The Allan variance is widely used in gyroscope performance analysis as an IEEE-approved standard analysis method. The results are shown in Table 2. On the one side, by using the improved method, the outputs of the MEMS gyroscope that ranged from −40 to 60 °C were reduced, and the temperature drift was greatly reduced. For example, the factor of Q was reduced from 1.2419 × 10⁻⁴ to 1.0533 × 10⁻⁶, the factor of B was reduced from 0.0087 to 1.8772 × 10⁻⁴, and the factor of N was reduced from 2.0978 × 10⁻⁵ to 1.4985 × 10⁻⁶. On the other side, the outputs of the MEMS gyroscope that ranged from 60 to −40 °C were reduced. The factor of Q was reduced from 2.9808 × 10⁻⁴ to 2.4430 × 10⁻⁶, the factor of B was reduced from 0.0145 to 7.2426 × 10⁻⁴, and the factor of N was reduced from 4.5072 × 10⁻⁵ to 1.0523 × 10⁻⁵.

Table 2. The Allan variance of the compared experiment based on original signal and the improved method.

|                   | Original Signal | Denoising Signal |
|-------------------|-----------------|------------------|
|                  | −40–60 °C       | 60–40 °C         | −40–60 °C       | 60–40 °C       |
| Q (°)             | 1.2419 × 10⁻⁴   | 2.9808 × 10⁻⁴    | 1.0533 × 10⁻⁶   | 2.4430 × 10⁻⁶  |
| N (°/√ horas)     | 2.0978 × 10⁻⁵   | 4.5072 × 10⁻⁵    | 1.4985 × 10⁻⁶   | 1.0523 × 10⁻⁵  |
| B (°/h)           | 0.0087          | 0.0145           | 1.8772 × 10⁻⁴   | 7.2426 × 10⁻⁴  |

6. Conclusions

The detailed temperature error of the MEMS gyroscope was studied by proposing an improved VMD-ELM algorithm based on CNN-LSTM and PSO-SVM. Within the improved fusion method based on the temperature experiment and compared experiment, the output of the MEMS gyroscope went through a process of searching for temperature error, establishing a temperature error compensation model, and filtering. The main findings are as follows:

(1) The improved VMD-ELM algorithm based on CNN-LSTM and PSO-SVM combined VMD, CNN-LSTM, PSO-SVM, and ELM. The final output of the MEMS gyroscope greatly decreased compared to that of the Allan variance method, which indicated good feasibility and effectiveness of the algorithms based on the novel method.

(2) Using the improved method, the output of the MEMS gyroscope ranging from −40 to 60 °C reduced, and the temperature drift dramatically declined. For example, the factor of Q reduced from 1.2419 × 10⁻⁴ to 1.0533 × 10⁻⁶, the factor of B reduced from 0.0087 to 1.8772 × 10⁻⁴, and the factor of N reduced from 2.0978 × 10⁻⁵ to 1.4985 × 10⁻⁶. Furthermore, the output of the MEMS gyroscope ranging from 60 to −40 °C reduced. The factor of Q reduced from 2.9808 × 10⁻⁴ to 2.4430 × 10⁻⁶, the factor of B reduced from 0.0145 to 7.2426 × 10⁻⁴, and the factor of N reduced from 4.5072 × 10⁻⁵ to 1.0523 × 10⁻⁵.

(3) The experiments show that the method proposed in this study can greatly compensate the gyro output signal to obtain zero bias stability, zero bias instability, and angle random walking with a stable effect. The experimental results show that the compensation of the novel VMD-ELM algorithm based on CNN-LSTM and PSO-SVM significantly improved with obvious compensation effects to provide certain engineering application value.

Author Contributions: X.W. and H.C. conceived and designed the experiments; X.W. and H.C. performed the experiments; X.W. and H.C. analyzed the data; X.W. and H.C. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by National Key Research and Development Program of China (No. SQ2022YFB3200085), National Natural Science Foundation of China (No. 51705477), Technology Field Fund of Basic Strengthening Plan of China (2020JCJQJJ409, 2021-JCJQ-JJ-0315), National defense
References

1. Cui, M.; Huang, Y.; Wang, W.; Cao, H. MEMS Gyroscope Temperature Compensation Based on Drive Mode Vibration Characteristic Control. *Micromachines* 2019, 10, 248. [CrossRef] [PubMed]

2. Cui, J.; Yan, G.; Zhao, Q. Enhanced temperature stability of scale factor in MEMS gyroscope based on multi parameters fusion compensation method. *Measurement* 2019, 148, 106947. [CrossRef]

3. Yin, T.; Lin, Y.; Yang, H.; Wu, H. A Phase Self-Correction Method for Bias Temperature Drift Suppression of MEMS Gyroscopes. *J. Circuits Syst. Comput.* 2020, 29, 12. [CrossRef]

4. Bu, F.; Fan, B.; Xu, D.; Guo, S.; Zhao, H. Bandwidth and noise analysis of high-Q MEMS gyroscope under force rebalance closed-loop control. *J. Micromech. Microeng.* 2021, 31, 065002. [CrossRef]

5. Xia, D.; Yu, C.; Kong, L. The Development of Micromachined Gyroscope Structure and Circuitry Technology. *Sensors* 2014, 14, 1394–1473. [CrossRef] [PubMed]

6. Pistorio, F.; Saleem, M.M.; Somà, A. A Dual-Mass Resonant MEMS Gyroscope Design with Electrostatic Tuning for Frequency Mismatch Compensation. *Appl. Sci.* 2021, 11, 1129. [CrossRef]

7. Nesterenko, T.G.; Koleda, A.N.; Barbín, E.S.; Uchaikin, S.V. Temperature Error Compensation in Two-Component Micromechanical Gyroscope. *IEEE Trans. Compon. Packag. Manuf. Technol.* 2014, 4, 1598–1605. [CrossRef]

8. Guo, Z.; Fu, P.; Liu, D.; Huang, M. Design and FEM simulation for a novel resonant silicon MEMS gyroscope with temperature compensation function. *Microsyst. Technol.* 2017, 24, 1453–1459. [CrossRef]

9. Fu, Q.; Di, X.-P.; Chen, W.-P.; Yin, L.; Liu, X.-W. A temperature characteristic research and compensation design for micro-machined gyroscope. *Mod. Phys. Lett. B* 2017, 31, 1750064. [CrossRef] [PubMed]

10. Cao, H.; Li, H.; Shao, X.; Liu, Z.; Kou, Z.; Shan, Y.; Shi, Y.; Shen, C.; Liu, J. Sensing mode coupling analysis for dual-mass MEMS gyroscope and bandwidth expansion within wide-temperature range. *Mech. Syst. Signal Process.* 2018, 98, 448–464. [CrossRef]

11. Cao, H.; Liu, Y.; Zhang, Y.; Shao, X.; Gao, J.; Huang, K.; Shi, Y.; Tang, J.; Shen, C.; Liu, J. Design and Experiment of Dual-Mass MEMS Gyroscope Sense Closed System Based on Bipole Compensation Method. *IEEE Access* 2019, 7, 49111–49124. [CrossRef]

12. Cao, H.; Cao, H.; Xue, R.; Xue, R.; Cai, Q.; Cai, Q.; Gao, J.; Gao, J.; Zhao, R.; Zhao, R.; et al. Design and Experiment for Dual-Mass MEMS Gyroscope Sensing Closed-Loop System. *IEEE Access* 2020, 8, 48074–48087. [CrossRef]

13. Wang, Y.; Liu, X.; Zhang, Y.; Fu, Q. A digital output MEMS DRG on-chip temperature compensation method based on virtual sensor. *Mod. Phys. Lett. B* 2020, 34, 2050422. [CrossRef]

14. Chong, S.; Song, R.; Li, J.; Zhang, M.; Tang, J.; Shi, B.; Liu, J.; Cao, H. Temperature drift modeling of MEMS gyroscope based on genetic-Elman neural network. *Mech. Syst. Signal Process.* 2016, 72, 897–905. [CrossRef]

15. Cao, H.; Zhang, Y.; Shen, C.; Liu, Y.; Wang, X. Temperature Energy Influence Compensation for MEMS Vibration Gyroscope Based on RBF NN-GA-KF Method. *Shock Vib.* 2018, 2018, 2830868. [CrossRef]

16. Shen, C.; Li, J.; Zhang, X.; Tang, J.; Cao, H.; Liu, J. Multi-scale parallel temperature error processing for dual-mass MEMS gyroscope. *Sens. Actuators A Phys.* 2016, 245, 160–168. [CrossRef]

17. Wu, Y.; Shen, C.; Cao, H.; Che, X.; Wu, Y.C.; Shen, C.; Cao, H.L.; Chen, X. Improved Morphological Filter Based on Variational Mode Decomposition for MEMS Gyroscope De-Noising. *Micromachines* 2018, 9, 246. [CrossRef] [PubMed]

18. Shi, H.; Hu, S.; Zhang, J. LSTM based prediction algorithm and abnormal change detection for temperature in aerospace gyroscope shell. *Int. J. Intell. Comput. Cybern.* 2019, 12, 274–291. [CrossRef]

19. Ma, T.; Li, Z.; Cao, H.; Shen, C.; Wang, Z. A Parallel Denoising Model for Dual-Mass MEMS Gyroscope Based on PE-ITD and SA-ELM. *IEEE Access* 2019, 7, 169979–169991. [CrossRef]

20. Chang, L.; Cao, H.; Shen, C. Dual-Mass MEMS Gyroscope Parallel Denoising and Temperature Compensation Processing Based on WLMP and CS-SVR. *Micromachines* 2020, 11, 586. [CrossRef] [PubMed]

21. Gu, H.; Zhao, B.; Zhou, H.; Liu, X.; Su, W. MEMS Gyroscope Bias Drift Self-Calibration Based on Noise-Suppressed Mode Reversal. *Micromachines* 2019, 10, 823. [CrossRef] [PubMed]

22. Song, J.; Shi, Z.; Du, B.; Han, L.; Wang, H.; Wang, Z. MEMS gyroscope wavelet de-noising method based on redundancy and sparse representation. *Microelectron. Eng.* 2019, 217, 111112. [CrossRef]
23. Ding, M.K.; Shi, Z.; Du, B.; Wang, H.; Han, L. A signal de-noising method for a MEMS gyroscope based on improved VMD-WTD. *Meas. Sci. Technol.* **2021**, *32*, 095112. [CrossRef]

24. Zhang, R.; Xu, B.; Wei, Q.; Yang, T.; Zhao, W.; Zhang, P. Serial-Parallel Estimation Model-Based Sliding Mode Control of MEMS Gyroscopes. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *51*, 7764–7775. [CrossRef]

25. Wang, P.; Li, G.; Gao, Y. A compensation method for gyroscope random drift based on unscented Kalman filter and support vector regression optimized by adaptive beetle antennae search algorithm. *Appl. Intell.* **2022**, *1*–16. [CrossRef]

26. Huang, F.; Wang, Z.; Xing, L.; Gao, C. A MEMS IMU Gyroscope Calibration Method Based on Deep Learning. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–16. [CrossRef]

27. Huang, F.; Wang, Z.; Xing, L.; Gao, C. A MEMS IMU Gyroscope Calibration Method Based on Deep Learning. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–16. [CrossRef]

28. Cao, H.; Li, H.; Kou, Z.; Shi, Y.; Tang, J.; Ma, Z.; Shen, C.; Liu, J. Optimization and Experimentation of Dual-Mass MEMS Gyroscope Quadrature Error Correction Methods. *Sensors* **2016**, *16*, 71. [CrossRef]

29. Cao, H.; Cai, Q.; Zhang, Y.; Shen, C.; Shi, Y.; Liu, J. Design, Fabrication, and Experiment of a Decoupled Multi-Frame Vibration MEMS Gyroscope. *IEEE Sens. J.* **2021**, *21*, 19815–19824. [CrossRef]

30. Ma, T.; Cao, H.; Shen, C. A Temperature Error Parallel Processing Model for MEMS Gyroscope based on a Novel Fusion Algorithm. *Electronics* **2020**, *9*, 499. [CrossRef]

31. Li, Z.; Gu, Y.; Yang, J.; Cao, H.; Wang, G. A Noise Reduction Method for Four-Mass Vibration MEMS Gyroscope Based on ILMD and PTTFPF. *Micromachines* **2022**, *13*, 1807. [CrossRef] [PubMed]