Article

Use of Long Short-Term Memory for Remaining Useful Life and Degradation Assessment Prediction of Dental Air Turbine Handpiece in Milling Process

Yi-Cheng Huang *[ID] and Yu-Hsien Chen

Department of Mechatronics Engineering, National Changhua University of Education, Changhua 50074, Taiwan; hjnm4512159@gmail.com
* Correspondence: ychuang@cc.ncue.edu.tw

Abstract: The complexity of the internal components of dental air turbine handpieces has been increasing over time. To make operations reliable and ensure patients' safety, this study established long short-term memory (LSTM) prediction models with the functions of learning, storing, and transmitting memory for monitoring the health and degradation of dental air turbine handpieces. A handpiece was used to cut a glass porcelain block back and forth. An accelerometer was used to obtain vibration signals during the free running of the handpiece to identify the characteristic frequency of these vibrations in the frequency domain. This information was used to establish a health index (HI) for developing prediction models. The many-to-one and many-to-many LSTM frameworks were used for machine learning to establish prediction models for the HI and degradation trajectory. The results indicate that, in terms of HI predicted for the testing dataset, the mean square error of the many-to-one LSTM framework was lower than that that of a logistic regression model, which did not have a memory framework. Nevertheless, high accuracies were achieved with both of the two aforementioned approaches. In general, the degradation trajectory prediction model could accurately predict the degradation trend of the dental handpiece; thus, this model can be a useful tool for predicting the degradation trajectory of real dental handpieces in the future.

Keywords: dental air turbine handpiece; long short-term memory; logistic regression; remaining useful life

1. Introduction

Dental air turbine handpieces are the most common medical equipment used by dentists for treating teeth. The main mechanical mechanism of the aforementioned equipment involves the introduction of high-pressure air into the head to rotate microturbine blades for driving bearings to rotate a tooth drill. Before or during treatment, a dentist cannot identify the state of health (SOH) of and the presence of any damage in the internal components of a dental handpiece. Because a dental handpiece has a high rotation speed, its internal bearing gradually degrades and becomes damaged after it has been used for a long time. Consequently, ceramic balls cannot operate stably along the designated track, which results in increased internal friction in the shell of the machine head and in increased temperature. Heat is transferred to the bur through the rotor, which increases the temperature of the bur. When the temperature of the bur at the milling contact surface is higher than 42.5 °C [1], it may cause irreversible damage to teeth. When the temperature reaches 52 °C, it may cause pulp necrosis. The vibrations of the bur make the patient uncomfortable and cause the tooth to have an uneven surface.

The authors of [2] used an accelerometer and a laser Doppler vibrometer to measure vibration signals during the operation of a dental handpiece. By comparing the spectrograms of the handpiece without loading (free running) and with loading (milling teeth), the aforementioned authors discovered that the vibration frequency decreased during

milling. The two measurement results did not exhibit significant differences. The authors of [3,4] used different operating conditions of dental air turbine handpieces to generate an experimental group and a control group. Dental handpieces were used according to pre-established testing procedures until their bearing failed, through which they analyzed factors affecting the service life of the bearing. The authors of [4] mentioned that the failure of the ball bearing was typically caused by faults in the cage (the nonmetal part of the bearing). When the bearing fails, the cage experiences substantial damage due to the running of the instrument under load and the occurrence of corrosion in high-pressure sterilization. Although it cannot be determined whether the wear and fracture of the cage directly leads to bearing faults, damage to the cage unavoidably hinders the stable operation of the ball in the raceways, which is sufficient to cause damage to the ball and raceways under relatively few rotations. Therefore, frictional heat is generated, which causes the formation of a heat-affected zone. When a dentist uses the handpiece, this heat is unavoidably conducted to the tooth drill.

Therefore, establishing a prediction model for the health index (HI) of dental handpieces is critical. Such a model can be used to monitor the SOH of a dental handpiece and predict its remaining useful life (RUL) in real time, which can allow dentists to identify the timing of potential equipment faults on the basis of comprehensive factors such as internal bearing damage, cage fracture, or loss of original dynamic balance by the rotor. Moreover, a dentist can use the aforementioned information to determine the time for replacing parts of the dental handpiece, such as the bearing, in the next usage cycle so that the handpiece does not suddenly malfunction during treatment and injure the patient.

In [5], the diagnostic results of a free-running of an air turbine dental handpiece (ATDH) with three rotor statuses by applying fast Fourier transform (FFT), Hilbert–Huang transform (HHT), and multiscale entropy (MSE) processes was studied. Our previously proposed method was tested under conditions of additional axial preload on the rotor and ceramic bearings with a damaged outer race supporting the rotor. Using a laser-Doppler vibrometer, condenser microphone, and portable MEMS system microphone were used to acquire the signals when the ATDH rotor features were changed. The results show that changes in preload or malfunctioning ball bearings can be discriminated and abstracted using FFT and HHT to analyze the vibration frequencies with three different sensing devices. However, no study has investigated the RUL of dental handpieces. Therefore, the current study used a long short-term memory (LSTM) network to establish an HI prediction model. Because an LSTM network can transfer memory and learn time-series data, this network is suitable for fault prediction and prognostic and health management, such as monitoring the system status and performing predictive maintenance.

Several studies have used different methods to construct an RUL system. The authors of [6,7] predicted the RUL of the bearing by using the moving-average cross-correlation based on power spectral density, a deep autoencoder, and a deep neural network. The authors of [8] used different RUL models for friction systems under different operation conditions. By using the signal trend and predefined threshold values, the aforementioned authors identified which model to use and estimated the RUL, thereby increasing the accuracy of the estimation. With advances in machine learning, studies have applied logistic regression (LR) to assess the degradation of machines. In [9], the wavelet packet decomposition technique was used to obtain features from data, such as current and vibration data, select critical features, and enter these features into the LR model for assessing machine performance and identifying possible failure models. With increases in the calculation speed of computers, deep learning has been applied practically in fields such as medicine and engineering. Numerous studies have adopted deep learning in prognostics and health management, including fault diagnosis and RUL prediction. In [10,11], the SOH and RUL of Li batteries were assessed. In [10], features were obtained from the incremental capacity curve to analyze the correlation between the features and the battery capacity. This analysis was conducted to select features with high correlations as the input for artificial neural network (ANN) assessment models. Moreover, the percentages of the

current capacity and initial capacity were used as standards to train assessment models to determine the SOH of lithium-ion batteries. Finally, the testing results of the validation dataset indicated that the ANN assessment models had high accuracy. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), which are two types of deep learning algorithms, have been widely used in various fields. CNNs are commonly used for handling classification problems such as image recognition and bearing fault diagnosis [12]. By contrast, RNNs are commonly used for handling time series problems. Because RNNs use the output from the hidden layer for the previous moment as the input of the hidden layer for the next moment, the output of each moment is associated with the input of the previous moment. Thus, RNNs can learn from and memorize time series. However, RNNs face the exploding gradient and vanishing gradient problems [13], rendering them difficult to train. LSTM networks can solve the vanishing gradient problem because, unlike in RNNs, in which the memory of each instance is replaced by a new memory, the previous memory is added to the new memory in LSTM networks. In [14], an LSTM model was used to predict oil prices. In [15,16], LSTM models were used in time-series-related research, such as RUL assessment. In [16], RUL indicators of turbofan engines were generated using LSTM models. Data obtained from sensors that exhibited relatively stable changes with the degradation of the engine were used in the aforementioned study. These data were subjected to processing procedures, such as filtration and normalization. Compared with linear decreasing, the physical signals obtained from the sensors could more accurately reflect the actual degradation of engines. Because the real SOH and RUL of a system are generally unknown, a piece-wise linear decreasing function was used in [17] to determine the RUL. For a healthy system, degradation is not obvious; therefore, the authors of [17] assumed that the RUL is initially constant and then decreases linearly. This study used an accelerometer to obtain the vibration signals of dental handpieces for analyzing the signal trends and features generated as the rotor of the dental handpieces gradually degraded with time. The identified trends and features were used as a critical degradation index to develop an HI model for assessing the SOH and RUL of dental handpieces.

2. Method

2.1. Many to One and Many to Many LSTM Structures for RUL and Degradation Assessment

An LSTM possesses a memory structure that contains memory cells. It adds and memorizes information as the time series progresses, thereby solving the vanishing gradient problem. Figure 1 illustrates the basic structure of an LSTM network. The cell state can be used to store and transmit memory; thus, the information in this state can only be written or deleted. Without external influence, the aforementioned information remains unchanged. The parameter x_t represents the input data at time t, and h_{t-1} is the hidden state at time $t-1$. The cell state at time $t-1$ is denoted as C_{t-1}, which is modified to the present cell state C_t in the hidden layer at time t.
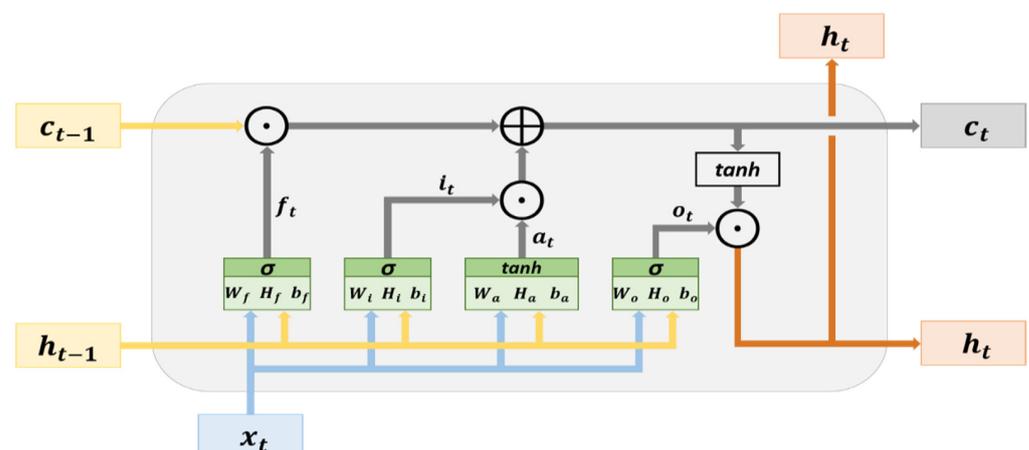

Figure 1. Architecture of LSTM memory cells.

The hidden layer of an LSTM network contains an input node (a_t) and three controlled gates (f_t, i_t, and o_t). The variables a_t, f_t, i_t, and o_t are calculated using (1)–(4), respectively. The input node a_t is used for updating the cell state, whereas the controlled gates are used to determine whether to allow information to pass through them. The controlled gates are the forget gate, input gate, and output gate. The forget gate (f_t) determines which cell states' (c_{t-1}) information may pass through it. The input gate (i_t) determines which input nodes' information (a_t) may pass through it. The vectors (information) passing through the input gate are used for updating the cell state and are subjected to element-wise addition with the vectors (information) passing through the forget gate to generate the cell state at time t (c_t). The calculation in the aforementioned process is expressed in (5). The output gate determines which cell states' (c_t) information may pass through it. The vectors (information) passing through the output gate are in the hidden state at time t (h_t), and they are the output vectors of the current hidden layer. The calculation method for h_t is presented in (6). In addition, the cell state and hidden state obtained at time t, namely (c_t) and (h_t), respectively, are transmitted to the hidden layer at time ($t + 1$). This process that progresses with the time series is used for the transmission and learning of memory.

$$a_t = tanh(W_a x_t + H_a h_{t-1} + b_a) \tag{1}$$

$$f_t = \sigma\left(W_f x_t + H_f h_{t-1} + b_f\right) \tag{2}$$

$$i_t = \sigma(W_i x_t + H_i h_{t-1} + b_i) \tag{3}$$

$$o_t = \sigma(W_o x_t + H_o h_{t-1} + b_o) \tag{4}$$

$$c_t = (f_t \odot c_{t-1}) \oplus (i_t \odot a_t) \tag{5}$$

$$h_t = o_t \odot tanh(c_t) \tag{6}$$

where W and H represent the weight, b denotes the bias, and $\oplus$ is the symbol for element-wise addition, $\odot$ is the symbol for element-wise multiplication, *tanh* denotes the hyperbolic tangent, and σ represents the sigmoid function. The parameters *tanh* and σ represent activation functions.

The current study adopted the many-to-one and many-to-many LSTM structures. These structures were used to develop different prediction models. When the current state was known, a model based on the many-to-one LSTM structure was used to predict the SOH of the next unknown cycle for the early assessment of the SOH of a dental handpiece. Consider the example of ten time steps. SOH data of the current known cycle and the nine cycles before it, namelexample of ten time steps. SOH data of the current known cycle and the nine cycles before it, namely $[x_{(t-9)}, x_{(t-8)}, x_{(t-7)}, \ldots, x_{(t)}]$, are entered into the LSTM model, which then outputs the SOH of the next unknown cycle, namely $[y'_{(t+1)}]$. Figure 2 illustrates an LSTM model with the many-to-one structure.
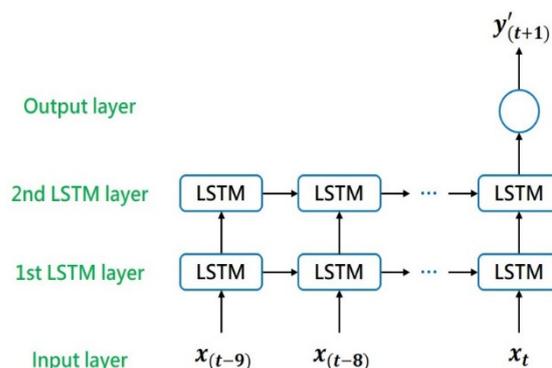

Figure 2. LSTM model with the many-to-one structure.

The many-to-many structure is used to predict the degradation trajectory of dental handpieces. Consider the example of ten time steps. The SOHs of the current known cycle and the nine cycles before it, namely $[y_{(t-9)}, y_{(t-8)}, y_{(t-7)}, \ldots, y_{(t)}]$, are input into the LSTM model, which then output the SOHs of the next ten unknown cycles, namely $[y'_{(t+1)}, y'_{(t+2)}, y'_{(t+3)}, \ldots, y'_{(t+10)}]$. Subsequently, the output results are used as the model input. The aforementioned model then outputs the SOHs of the next ten unknown cycles, namely $[y'_{(t+11)}, y'_{(t+12)}, y'_{(t+13)}, \ldots, y'_{(t+20)}]$. This method is repeated for iteration to predict the degradation trajectory of a dental handpiece. Figure 3 shows an LSTM model with the many-to-many structure. In the many-to-one structure, in each time step, the model can accept multidimensional input data. By contrast, in the many-to-many structure, in each time step, input data must be unidimensional to enable the model to continue iterating.
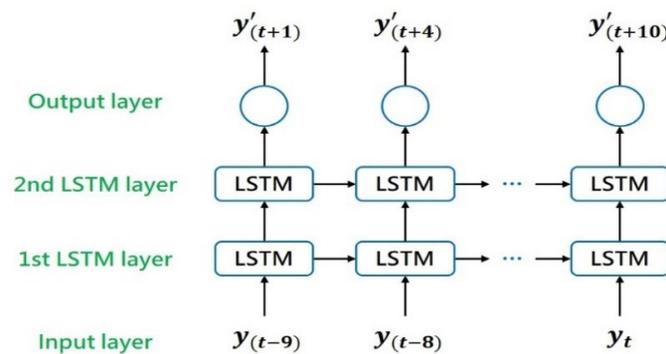

Figure 3. LSTM model with the many-to-many structure.

2.2. Logistic Regression Prediction Model for Many to One Structure

In Equation (7), the sigmoid function is used for predicting the output of LR with the inputs features $x_1 - x_n$ in Equation (8) using model parameters $\theta_0 - \theta_n$. A conventional run-to-failure LR sets y as a constant to 1 for the initial healthy status, to 0 for failure, and to 0–1 for the rest of the work asset's lifetime. To build the RUL model accurately with a LR, the difference between the predicted $h_\theta(x)$ and y is minimized. Therefore, the model parameters for each $\theta_0 - \theta_n$ are trained based on Equation (7) to predict the actual CV, y, of the machine's health. The model parameters for $\theta_0 - \theta_n$ can be derived using the Bernoulli distribution, maximum likelihood estimation (MLE), a loss function, and the method of gradient descent.

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}} \tag{7}$$

$$\theta^T x = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \ldots + \theta_n x_n \tag{8}$$

The probability by Bernoulli distribution is given by Equation (9).

$$p(y|x) = h_\theta(x)^y (1 - h_\theta(x))^{1-y} = \begin{cases} h_\theta(x) & if\ y = 1 \\ 1 - h_\theta(x) & if\ y = 0 \end{cases} \tag{9}$$

The MLE function is applied by multiplying each probability in Equation (9) after the most likely real outputs y are obtained from the inputs $X_1 - X_m$ and the model parameters $\theta_0 - \theta_m$. Thus, Equation (10) provides a likelihood estimation function based on m known samples and the loss function in Equation (11). Iterations of the gradient descent method with the learning rate (α) to find the value of each model parameter θ_j are performed according to Equations (12) and (13).

$$L(\theta) = \prod_{i=1}^{m} p(y^{(i)}|x^{(i)}; \theta) = \prod_{i=1}^{m} h_\theta(x^{(i)})^{y^{(i)}} \left(1 - h_\theta\left(x^{(i)}\right)\right)^{1-y^{(i)}} \tag{10}$$

$$J(\theta) = -\log L(\theta) = \sum_{i=1}^{m} \left(\log\left(1 + exp^{\theta^T x^{(i)}}\right) - y^{(i)}\theta^T x^{(i)} \right) \tag{11}$$

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{\partial \sum_{i=1}^{m} \left(\log\left(1 + exp^{\theta^T x^{(i)}}\right) - y^{(i)}\theta^T x^{(i)} \right)}{\partial \theta_j} = \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})x_j^{(i)} \tag{12}$$

$$\theta_j = \theta_j - \alpha \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})x_j^{(i)} \tag{13}$$

The LR structure is used to predict the SOH of the next unknown cycle on the basis of the current known state; thus, the LR structure can achieve the early assessment of the SOH of a dental handpiece. Because the LR structure does not possess the cell state of the LSTM network for storing and transmitting memories, the LR structure directly transmits the known state into the model. Consider the example of ten time steps. SOH data of the current state and the nine cycles before it, namely $[y_{(t-9)}, y_{(t-8)}, y_{(t-7)}, \dots, y_{(t)}]$, are input simultaneously into the LR model, which then outputs the SOH of the next unknown cycle, namely $[y'_{(t+1)}]$. The structure of the LR model is displayed in Figure 4. The current study compared the LSTM model, which possesses the memory storage and transmission functions, with the LR model, which does not possess these functions, and analyzed their prediction results.
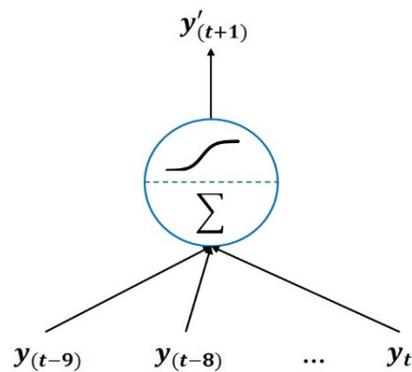

Figure 4. Illustration for LR model.

2.3. Kalman Filter

A Kalman filter (KF) combines the least square method and state-space representation in the recursive solution of linear filtering for a dynamic system. The KF is one of the optimal recursive algorithms in data processing, particularly in the field of navigation, communication, and satellite and flight control, where high tolerance to noise is crucial. Contrary to the ordinary concept of low-pass, high-pass, and band-pass filters, the KF is an estimation and prediction approach based on a probability density function. The KF, with the optimal recursive data processing algorithm, can estimate the present state by estimating the previous step and observing the current state. Therefore, without adopting previous observations and estimation data, less memory space is used, and a shorter system response time can be obtained. Furthermore, the KF outperforms its low- and high-pass counterparts in handling probabilistic noises.

During execution, operations of the KF can be divided into two parts, namely prediction and correction. In the recursive process, the following notations are employed [18]:

$\widetilde{x}_k$: posteriori estimate vector at time step k
A_k: state transition matrix at time step k, $A \in R^{n \times n}$
B_k: input control matrix at time step k, $B \in R^{n \times 1}$
z_k: measurement vector at time step k
H_k: measurement matrix at time step k
v_k: measurement noise at time step k
Q_k: process noise covariance matrix

R_k: measurement noise covariance matrix

Q_k and R_k are the influences of the external disturbance noise and measurement noise, respectively, on the system. The recursive process of the KF can be explained as follows:

I. Prediction

State prediction:

$$\widetilde{x}_k^- = A_k \widetilde{x}_{k-1} + B_k u_{k-1} \tag{14}$$

Error covariance matrix prediction:

$$P_k^- = A_k P_{k-1} A_k^T + Q_k \tag{15}$$

II. Correction

State correction:

$$K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k \right)^{-1} \tag{16}$$

$$\widetilde{X}_k = \widetilde{x}_k^- + K_k (z_k - H_k \widetilde{x}_k^-) \tag{17}$$

$$P_k = (1 - K_k H_k) P_k^- \tag{18}$$

In the aforementioned procedure, estimation and prediction using the previous time step can be employed for estimating and predicting the subsequent time step; that is, the previous state is used to determine the subsequent state. In this study, the state transition matrix is performed as a single variable at a time for filtering.

3. Experiments

3.1. Experimental Setup and Milling

The experimental instruments consisted of a dental device platform and a dental air turbine handpiece (Figure 5). The dental air turbine handpiece (Tiger101-3T4, 300×10^3–360×10^3 rpm, Thunder, Tiger group, Taichung, Taiwan) contained a machine head, hand grip, and handpiece connector. The adopted carbide bur (FG 558) had a diameter of 1.0 mm and a total length of 19 mm. This study used an Alicat PC-Series proportional–integral–derivative (PID) single-valve pressure controller to control the drive air pressure (Figure 6). An RS-232 signal line was connected to the Flow Vision SC software program on a personal computer. On the computer, various parameters related to the pressure control valve were adjusted to determine the location of this valve and the input voltage. The input voltage was adjusted to control the set point. A triaxial desktop computer numerical control (CNC) machine (mini-CNC P RX 1510, Original Mind Co., Nagano, Japan) was used to simulate the motions of humans operating the dental handpiece. This machine was used to move the dental handpiece and a glass ceramic block to conduct a cutting experiment. The material used for cutting was glass ceramic tempered with IPS e.max CAD. A triaxial accelerometer (sensitivity: 1000 mV/g) was glued to a jig. The NI cDAQ-9174 and NI 9230 instruments were used to measure and capture the vibration signals of the handpiece during free running. The collected data were transferred to and stored in LabVIEW. The sampling frequency was set at 12.8 kHz. The overall experimental setup is illustrated in Figure 6. The cutting conditions were as follows: feed rate = 100 mm/min, air pressure = 50 psi, and cutting depth = 0.1 mm. The cutting process involved linear back-and-forth milling (up- and down-milling; Figure 7). The cutting width was 0.2 mm until the glass ceramic was cut to a depth of 0.1 mm, which was the cutting path of a cycle, as illustrated in Figure 8. After completing the cutting path of a cycle, the vibration signals of the dental handpiece during free running were obtained during the cycle. In this manner, free-running vibration signals were obtained for different cycles. In each cycle, a volume of approximately 17.87 mm^3 was removed. Each cycle lasted approximately 20 min. Excluding the factor of mill wear, a new tooth drill (based on several experimental experiences) could generally be used for 20 cycles. Then, according to the visual inspection results, the tooth drill was replaced.
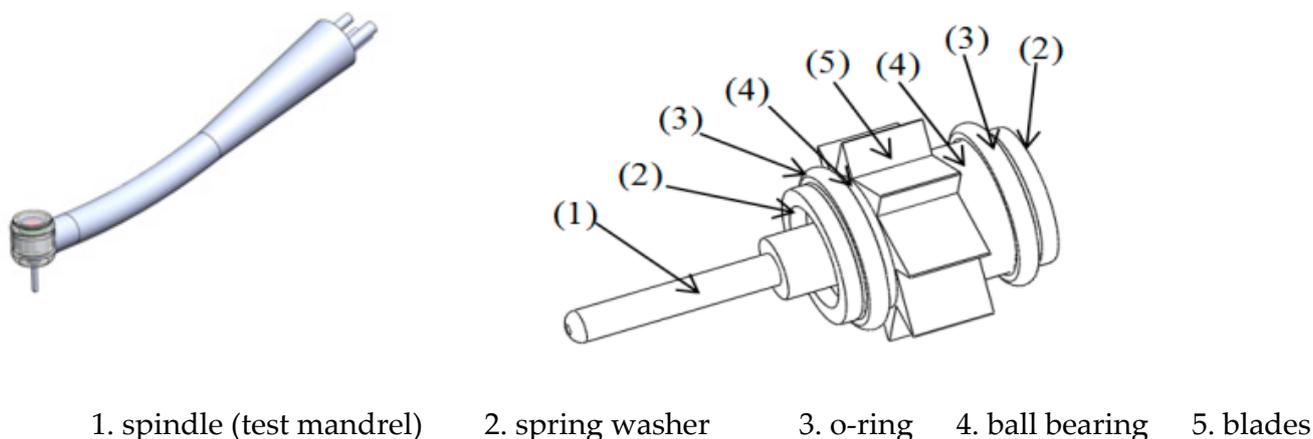
1. spindle (test mandrel)   2. spring washer   3. o-ring   4. ball bearing   5. blades

Figure 5. Schematic illustration and picture of in-house ATDH measurement platform.
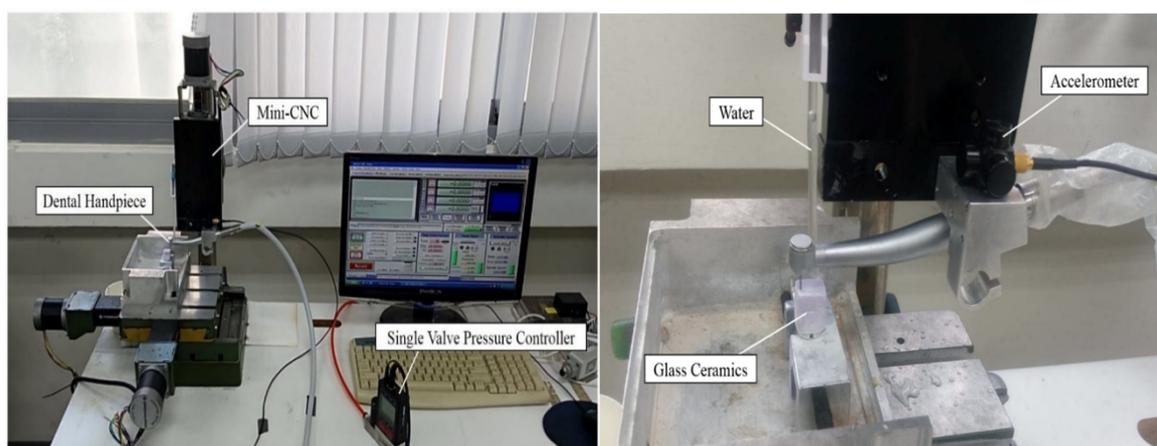

Figure 6. Triaxial CNC dental handpiece used in the cutting experiment.

Two new dental handpieces were used in the experiment. After dental handpiece No. 1 completed 250 cycles (5000 min) of the cutting experiment, it did not reflect clear degradation characteristics in the time- or frequency-domain analyses. Considering the reliability factors of handpiece production, the cutting experiment was continued. To quicken the degradation of the aforementioned dental handpiece, after approximately 300 cycles, the conditions of the cutting experiment were altered. The cutting feed rate was reduced to 30 mm/min; the air pressure was decreased to 45 psi; the cutting depth was increased from 0.1 to 0.2 mm; the cutting width was 0.3 mm; and the cutting process involved only down-milling. The reason why the cutting depth and width were increased was that incomplete cutting was observed in cycles 250–300 during the up-milling process. After confirming the cutting conditions and paths, for each cycle starting from the 300th

cycle, the removed volume increased to 35.74 mm^3. Moreover, the time required for the completion of each of these cycles was approximately 47 min. As the cutting experiment progressed, in some cycles, the dental handpiece exhibited the phenomenon of a sudden drop in rotation speed. Although the handpiece appeared to be stable, after approximately 375 cycles, the dental handpiece could not conduct more than 20 cycles of the cutting experiment after a new tooth drill was attached to it. A new tooth drill can typically be used for more than 20 cycles. Therefore, it could be inferred that the dental handpiece had degraded. The cutting experiment of dental handpiece No. 1 was stopped at this moment. This handpiece conducted 416 (250 + 166) cycles of the cutting experiment. The conditions of the cutting experiment for dental handpiece No. 2 were as follows: cutting feed rate = 100 mm/min, air pressure = 50 psi, and cutting depth = 0.1 mm. The cutting process involved linear back-and-forth milling (Figure 7). The cutting width was 0.2 mm until the glass ceramic block was cut to a depth of 0.1 mm. After 250 cycles of the cutting experiment, vibration signals with characteristic frequencies that were not observed under normal conditions gradually appeared; thus, the cutting experiment was terminated.
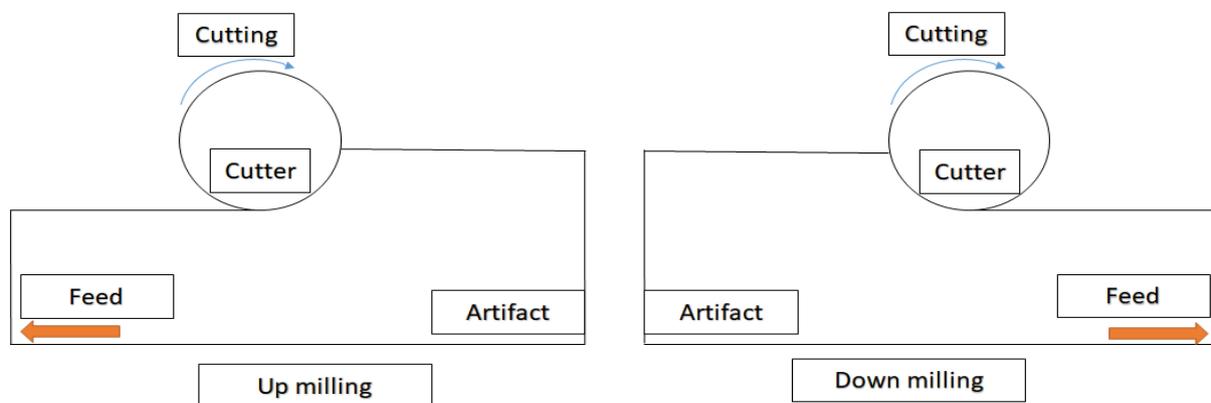

Figure 7. Schematics of up-milling (**left**) and down-milling (**right**).
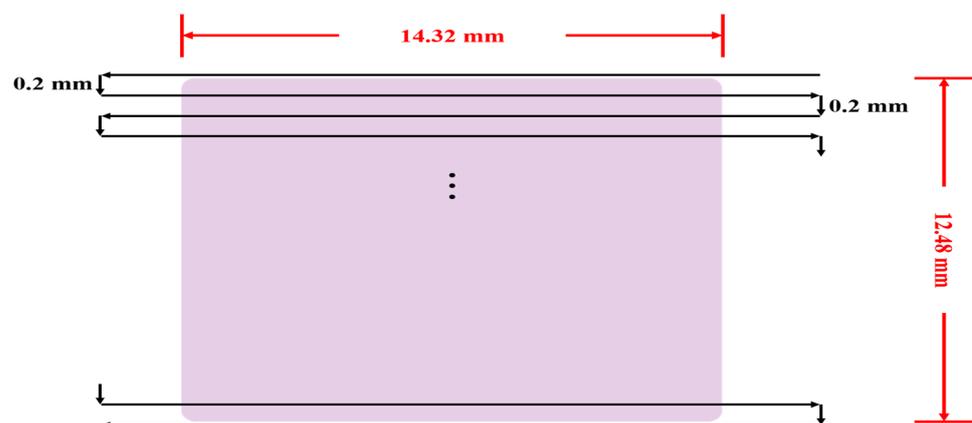

Figure 8. Schematic of the cutting path for one cycle.

3.2. Free-Running Vibration Signals

In [5], a noncontact laser Doppler vibrometer, a condenser microphone, and a mobile Micro Electro Mechanical System (MEMS) microphone were used to measure the vibration signals of a dental handpiece under free running. Under the no-loading (free running) condition, no significant differences were observed in the measurements of the aforementioned instruments. Under in vivo medical conditions, the mobile MEMS microphone was the most suitable instrument for the diagnosis of the health of the handpiece [5]. The authors of [2] used an accelerometer and a laser Doppler vibrometer to measure the vibration signals of a dental handpiece during its operation. The measurement results

obtained with these two instruments were plotted into the frequency spectrum. Under the no-loading (free running) condition, the measurement results of these two instruments did not exhibit substantial differences. Therefore, the current study used the FFT to transform the free-running vibration signals obtained from the accelerometer into frequency-domain signals (Figure 9). Dental handpiece No. 1 and dental handpiece No. 2 operated under the same conditions up to 250 cycles. After the 330th cycle, dental handpiece No. 1 began to exhibit problems such as an unstable rotation speed and a sudden decrease in the rotation speed. Moreover, the material removal was incomplete.

Figure 9. Main rotation speed frequencies of dental handpiece No. 2 (**left**) and dental handpiece No. 1 (**right**).

Except for the slight decreases in approximately the 150th cycle, basic frequency (BF) and ball pass frequency of the outer race of dental handpiece No. 2 exhibited no substantial differences throughout the 250 cycles. Figure 9 illustrates the FFTs of the 10th cycle (left) and 200th cycle (right). Figure 10 indicates that during the cutting experiment, the characteristic frequency (CF) of the vibration exhibited two additional subcharacteristic frequencies in each cycle (frequencies labeled 1 and 2 in the right part of Figure 10), and the vibration amplitude gradually increased with the number of cycles. The main frequency (in the left of Figure 10, 4 kHz) gradually decreased as the number of cycles increased. The handpiece components were gradually damaged under a fixed drive energy (50 psi); therefore, other vibration signals were transmitted.
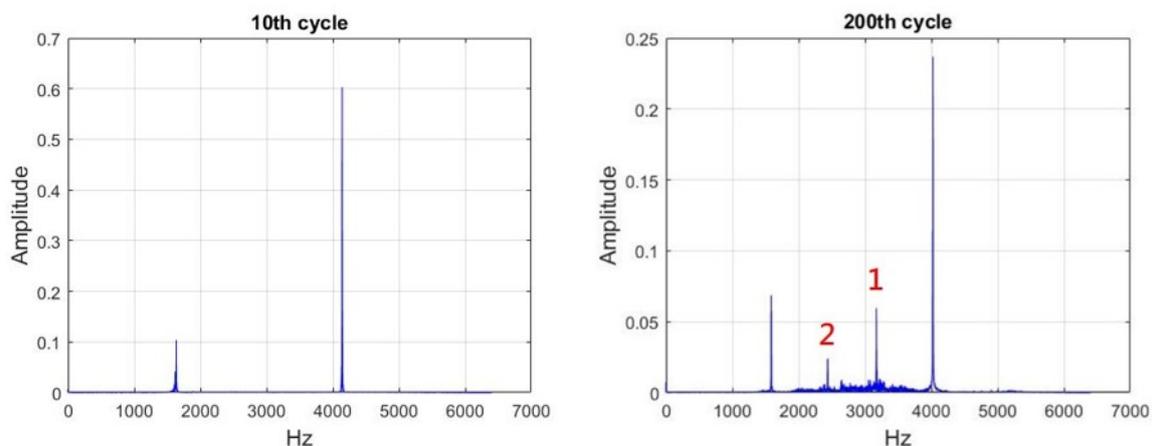

Figure 10. FFTs of the 10th cycle (**left**) and 200th cycle (**right**) for dental handpiece No. 2.

For the frequencies labeled 1 and 2 in Figure 10, CF(1) $\cong$ 3200 Hz and CF(2) $\cong$ 2500 Hz. These frequencies were repetitive and noticeable. Therefore, the cycles with an unstable frequency domain in Figure 11 were considered as outliers and removed to facilitate

subsequent analysis. A total of 17 cycles were removed. The remaining 233 cycles were used as experimental data to establish the HI and prediction models.
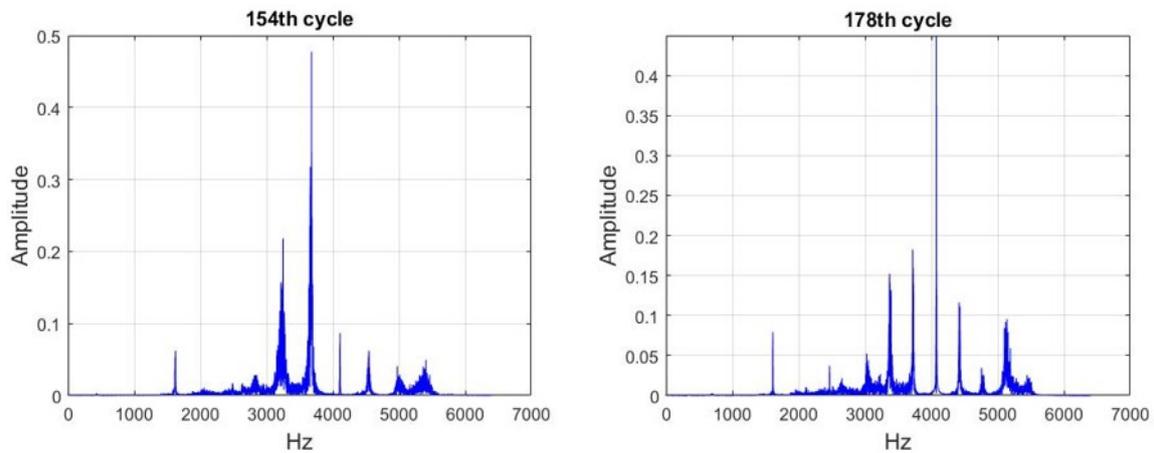

Figure 11. Cycles with unstable frequency domains (the 154th and 178th cycles).

Figure 12 illustrates the distribution of CF1, CF2, and the BF for each cycle. As shown in Figure 13 The results indicate that CF1 and CF2 were approximately 78% (r_1) and 60% (r_2) of the main frequency, respectively.
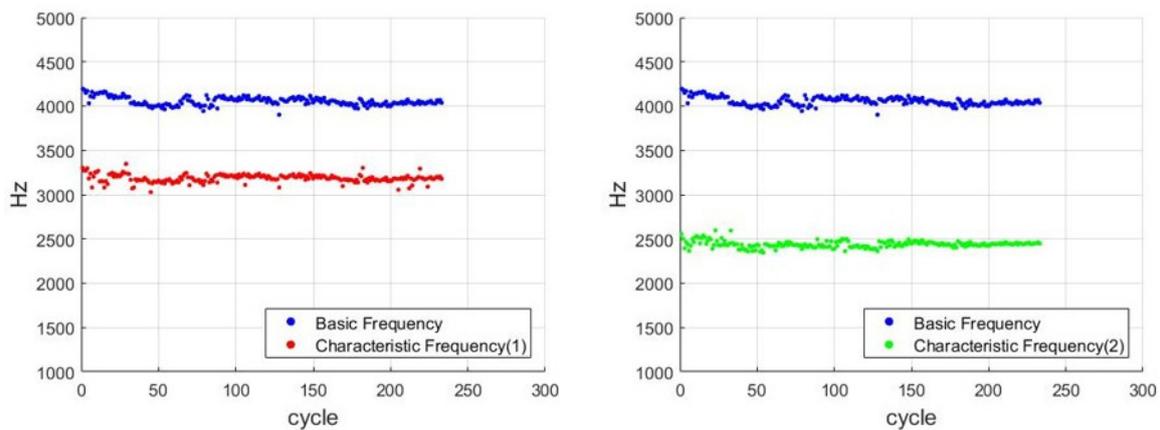

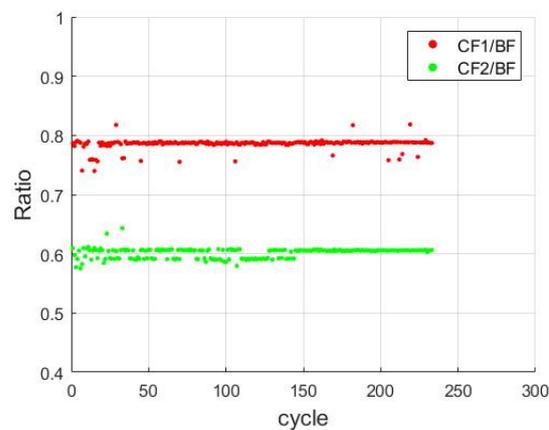**Figure 12.** Distributions of the BF, CF1, and CF2 for each cycle.

Figure 13. CF1/BF and CF2/BF ratios for each cycle.

To determine whether CF1 and CF2 changed with the progress of the cutting experiment, the trends of these CFs were analyzed by calculating the amplitude ratios of these CFs to the entire frequency domain. The relevant formula is expressed as follows:

$$R(r_j) = \frac{1}{S} \sum_{i=-150}^{150} C(BF \times r_j + i), \ j = 1, 2 \tag{19}$$

$$S = \sum_{k=1}^{6400} C(k) \tag{20}$$

where $C(x)$ is the amplitude of the x-section of the frequency distribution of 1–6400 Hz; x is the xth section; r denotes the ration; $r_1 = 0.78$ and $r_2 = 0.6$ for CF1 and CF2, respectively; S is the sum of the amplitude of the entire frequency domain; and R is the ratio of the sum of the CF amplitude to the amplitude sum of the entire frequency domain. The sampling frequency of the capture card adopted in this study was 12.8 kHz. After performing FFT, the study could analyze a frequency of up to 6.4 kHz. Therefore, this study calculated the sum of amplitudes of frequency up to 6.4 kHz. Because the air pressure tube providing air pressure to the dental air turbine handpieces was an elastic plastic tube, the provided air pressure was not completely consistent even under PID control, which resulted in some fluctuation in the CF. Therefore, amplitudes in the frequency ranges of CF1 $\pm$ 150 Hz and CF2 $\pm$ 150 Hz were summed and used as the CF amplitudes.

Figure 14 depicts the R values of CF1 and CF2 obtained for each cycle after using (19) and (20). The results indicate that as the experiment proceeded, $R_{(r1)}$ and $R_{(r2)}$ substantially increased. These parameters represented CF amplitudes relative to overall amplitudes. Under long-term milling, the amplitude of the CFs of the handpiece components increased; thus, as the number of cycles increased, the dental handpieces were gradually damaged, which resulted in the formation of CFs. Consequently, $R_{(r1)}$ and $R_{(r2)}$ are suitable for being used to establish the HI of the dental handpieces.
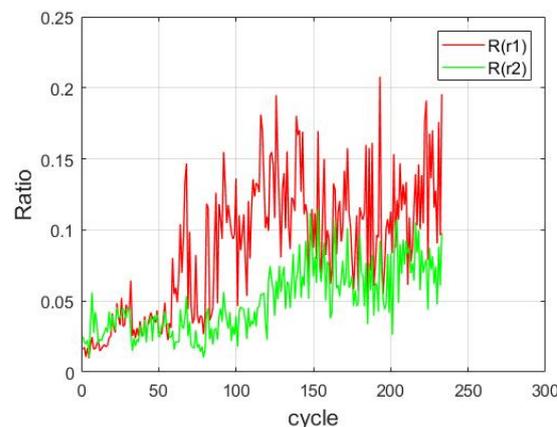

Figure 14. R values of CF1 and CF2 in relation to the amplitude of the entire frequency domain.

3.3. Building Health Index

The Kalman filter was applied to $R_{(r1)}$ and $R_{(r2)}$ to smooth out the curve in Figure 14 and obtain trend lines in Figure 15 with the coefficients of Q and R by 0.00001 and 0.0013, respectively. The trend lines were normalized to between 0 and 1. All the values were subtracted from the maximum value in the data, which resulted in a reversed trend line, which is represented by (21). In (21), x_i represents the data that required normalization, x_{max} is the maximum value in the data, x_{min} is the minimum value in the data, and X_n is the normalized data. The aforementioned normalization was conducted to generate an HI in which 1 represented the initial healthy state and 0 represented the most damaged state. The SOH of each cycle was between 0 and 1. Finally, because the two HIs generated from the R values were critical indices for the health of the dental handpieces, these HIs were

averaged in each cycle to obtain the final HI of the handpieces. Figure 15 illustrates the complete process for generating the HI.
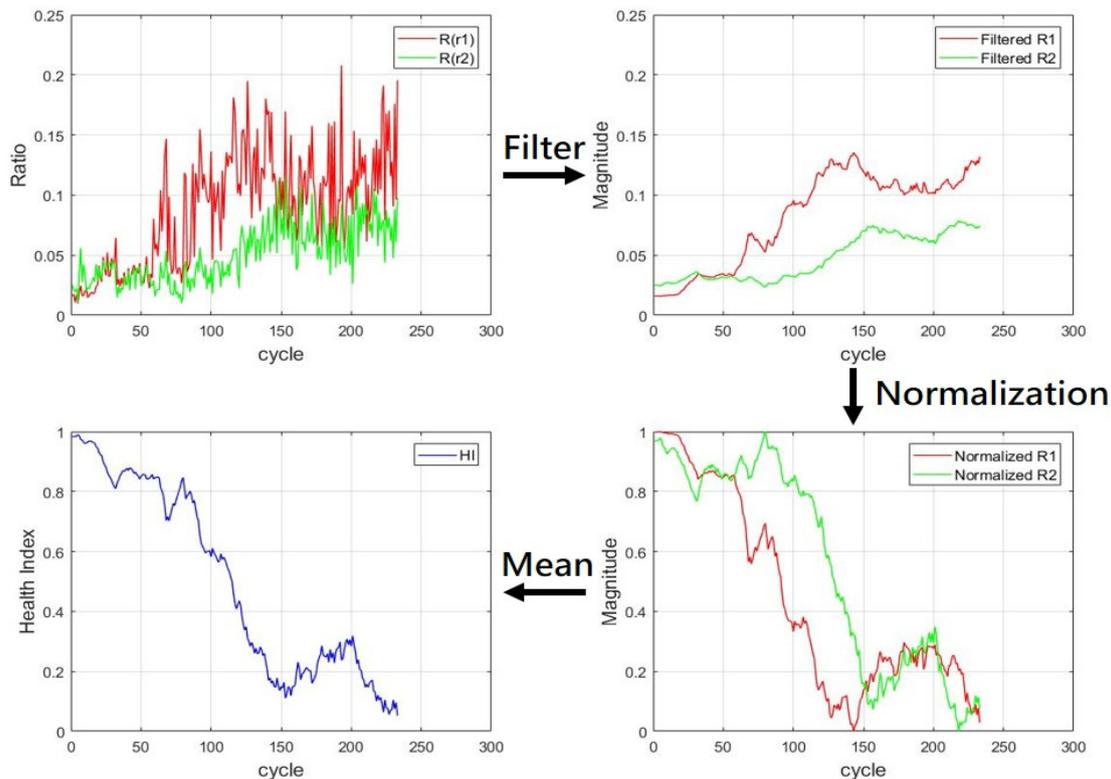
$$X_n = \frac{x_{max} - x_i}{x_{max} - x_{min}} \tag{21}$$

Figure 15. Procedure for generating the HI.

4. Remaining Useful Life Prediction and Degradation Assessment

4.1. RUL Based on LSTM Many-to-One Structure

To train and test the LSTM model, odd and even cycles were used as the training and testing datasets, respectively. Therefore, the training and testing datasets comprised 117 and 116 cycles, respectively. The data remained matched to the same cycles after the data segmentation. The model parameters were as follows: the number of time steps was 10, the number of epochs was 100, the loss function was the mean square error, and the number of LSTM hidden layers was 2.

An LSTM network with a many-to-one structure was used to establish an assessment model for the condition in which the state of a cycle is known to predict the SOH of the next unknown cycle. This study used the two normalized values as the model input. The two normalized values were from the definition of Equations (19) and (20) for R(r1) and R(r2), respectively. The filtered, normalized and array by two of $R_{(r1)}$ and $R_{(r2)}$ data are treated as $x_{(t)}$ for the LSTM's manipulation. Therefore, the input data were $[x_{(t-18)}, x_{(t-16)}, x_{(t-14)}, \ldots, x_{(t)}]$, the model output was $[y'_{(t+2)}]$, and the corresponded target HI was $[y_{(t+2)}]$. After the LSTM model was subjected to training with the training dataset, the mean square error gradually reduced and converged (Figure 16). Figure 17 illustrates the prediction results of the aforementioned model. This figure indicates that after training, the aforementioned LSTM model accurately predicted the SOH even with a dataset unknown to it.
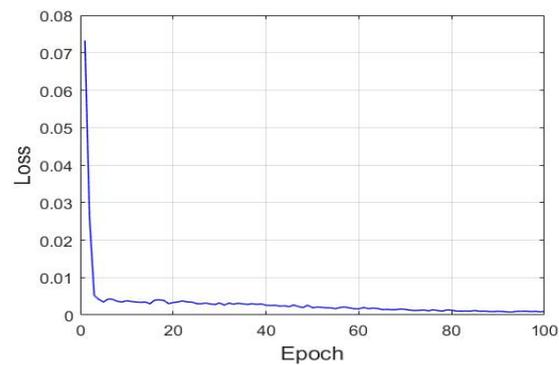
Figure 16. Gradual convergence of the loss function of the LSTM model with the many-to-one structure.
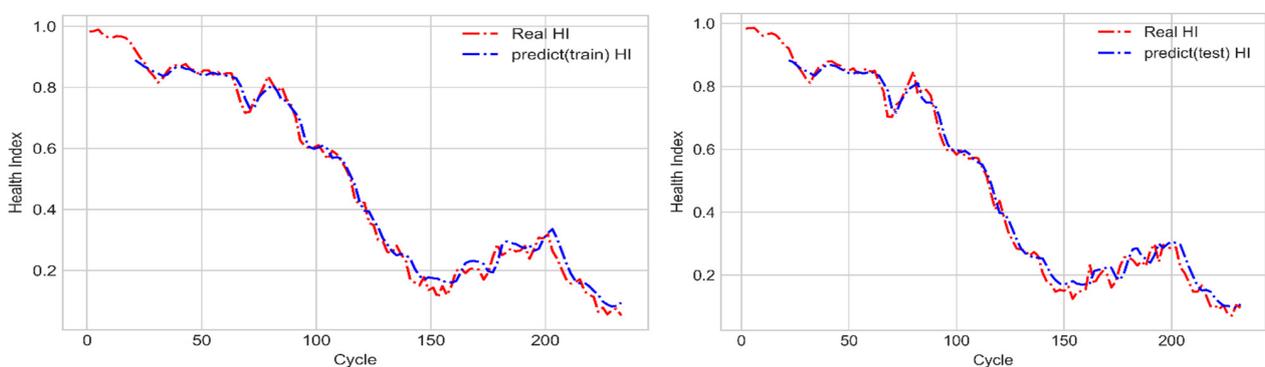

Figure 17. Prediction results obtained with the LSTM model with the many-to-one structure for the training dataset (**left**) and testing dataset (**right**).

4.2. LR Prediction Model

The LR prediction model was used to establish an assessment model for using known cycle state to predict the SOH of the next unknown cycle. The input data of this model were $[y_{(t-18)}, y_{(t-16)}, y_{(t-14)}, \ldots, y_{(t)}]$, the model output was $[y'_{(t+2)}]$, and the corresponding target HI was $[y_{(t+2)}]$. After conducting training with the training dataset, the prediction results of the LR model were compared with those of the LSTM model with the many-to-one structure (Figure 18). According to Figure 18, both these models had favorable accuracies. Therefore, the mean square errors of their prediction results were analyzed. As shown in Table 1, the mean square errors of the prediction results of the LSTM model with the many-to-one structure had lower errors than those of the LR model for both the training and testing data.
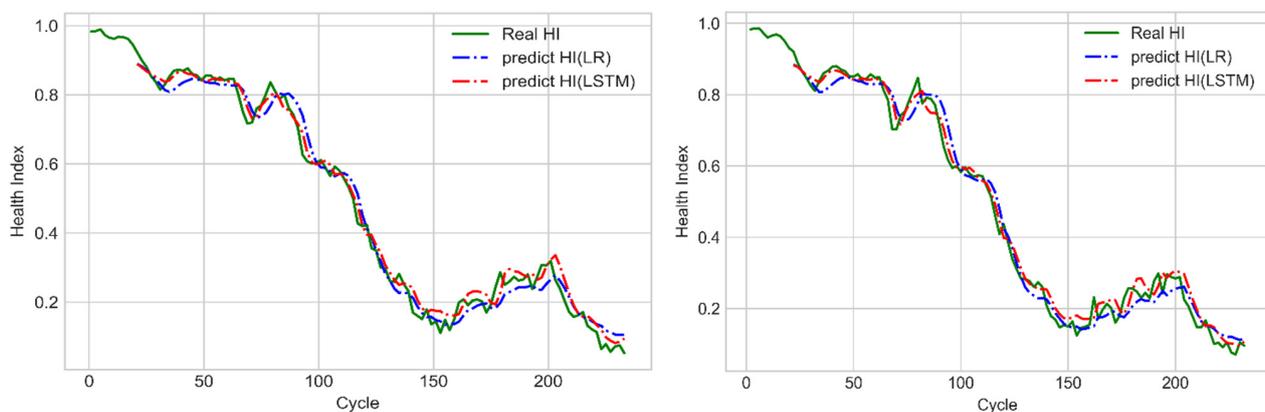

Figure 18. Prediction results of the LSTM model with the many-to-one structure and the LR model for the training dataset (**left**) and testing dataset (**right**).
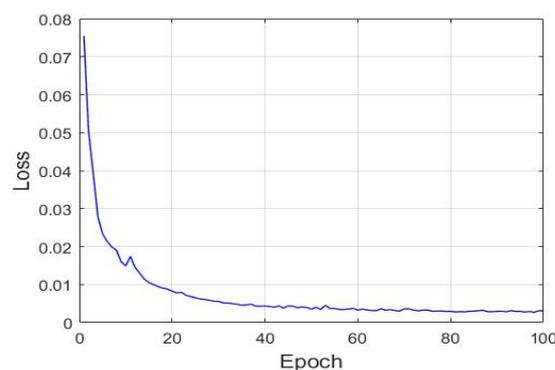
Table 1. Mean square errors of the prediction results of the LR model and the LSTM models with the many-to-one structure.

	LSTM	LR
Training dataset	0.000769	0.001359
Testing dataset	0.000820	0.001360

As previously mentioned, a conventional run-to-failure LR sets y as a constant to 1 for the initial healthy status, to 0 for failure, and to 0–1 for the rest of the work asset's lifetime. Here, two feature extractions (CF1 and CF2) were fed into the LR model, and resulted in good accuracy. However, the real SOH and RUL of a system were generally unknown in advance, and were changed with time. Raw data for time-forcasting in LR may not apply. In Section 4.1, the LSTM is able to work on the raw sensory data of machine's condition. The advantage of LSTM is to discover some unseen and hidden structure to improve the general model. The LSTM many-to-one for RUL prediction can do the monitoring job for RUL's prediction based on few milling process cycles. In Section 4.3, the LSTM's many-to-many structure will be introduced for the SOH of the next unknown cycles. It is used for the degradation assessment in practice. And it can be deployed whenever one wants to set a threshold for a diagnostic issue or performs a prognostic diagnosis.

4.3. LSTM Many-to-Many Structure

An LSTM model with the many-to-many structure was established for predicting the degradation trajectory of a dental handpiece. In the training stage, the SOH of the known cycle was used as the model input data. Therefore, the input data were $[y_{(t-18)}, y_{(t-16)}, y_{(t-14)}, \ldots, y_{(t)}]$, the model output was $[y'_{(t+2)}, y'_{(t+4)}, y'_{(t+6)}, \ldots, y'_{(t+20)}]$, and the corresponding target HI was $[y_{(t+2)}, y_{(t+4)}, y_{(t+6)}, \ldots, y_{(t+20)}]$. During the testing stage, the predicted output of the aforementioned model was used as the input for the next model prediction; thus, iterations were executed for generating the degradation trajectory. After training with the training dataset, the aforementioned LSTM model's mean square error gradually reduced and converged (Figure 19).

Figure 19. Gradual convergence of the loss function of the LSTM model with the many-to-many structure.

In the testing stage, 30, 60, and 90 known cycles of SOHs were input into the LSTM model with the many-to-many structure for predicting the subsequent degradation trajectory. The results obtained with the aforementioned model for the training dataset are displayed in Figure 20. A note is made here that the trajectories of known HI and real HI in Figure 20 were all from the same conducted experimental results. According to the two figures at the top of Figure 20, when the aforementioned LSTM model was used to predict the degradation trajectory, it reflected a gradually declining SOH in the middle cycles; thus, the aforementioned model achieved favorable prediction results with respect to the target HI. However, in the later cycles, the predicted SOH did not decline; thus, the degradation trajectory did not reflect the target HI well. Figure 21 shows the assessment of the testing

dataset by the LSTM model with the many-to-many structure. Although this model was not trained with the testing dataset, the predicted degradation trajectory was similar to that presented by the training dataset. Table 2 indicates that when the number of known cycles was 90, the prediction results for both the training and testing datasets exhibited the largest mean square errors. This result was obtained because the aforementioned model could not favorably predict the SOH in later cycles, which resulted in a large error. Additionally, the two datasets were subsets of the same dataset; therefore, the mean square error of the testing dataset was lower than that of the training dataset.
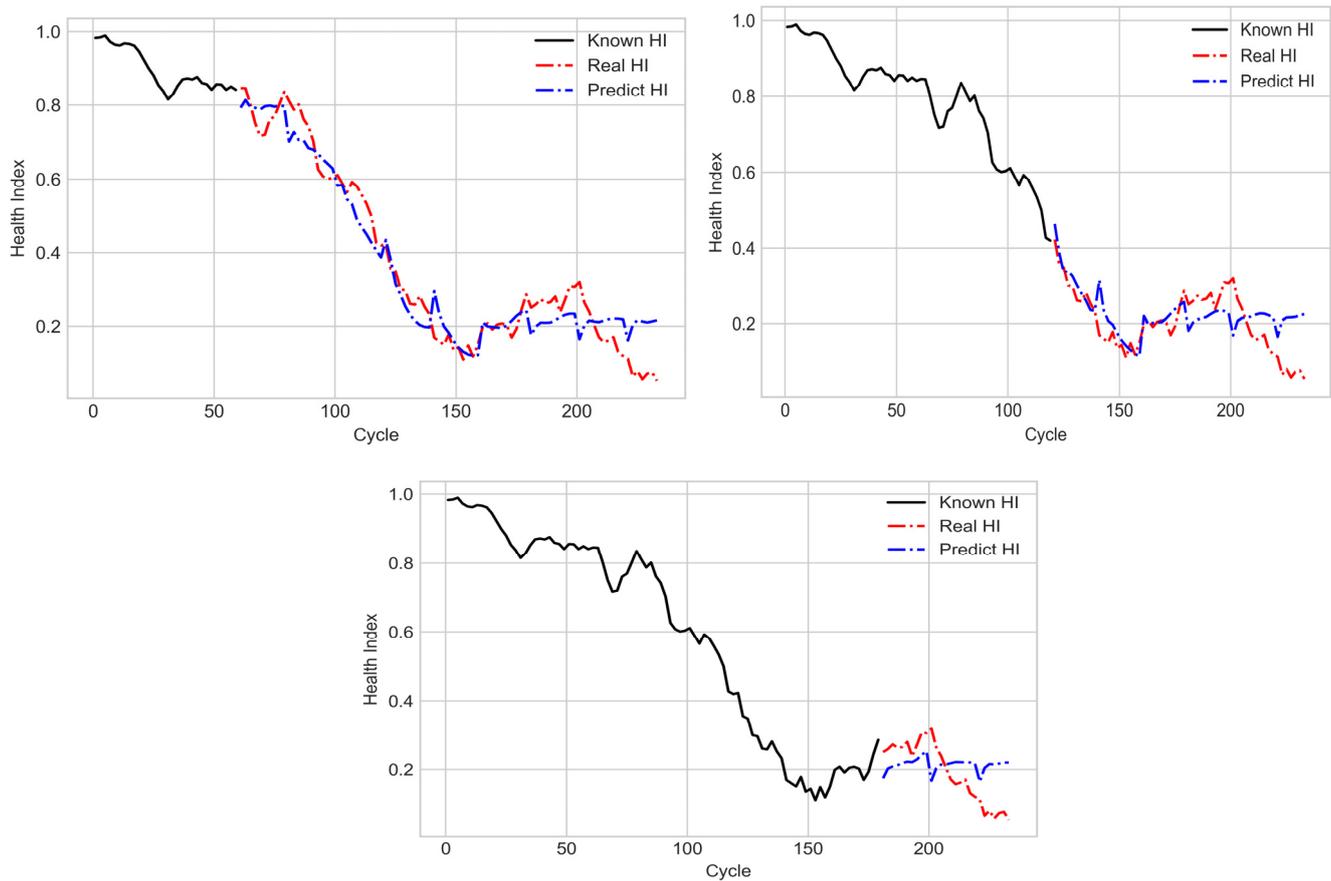

Figure 20. Predicted degradation trajectory for the training dataset under different numbers of known cycles.

Table 2. Mean square errors of the prediction results of the LSTM model with the many-to-many structure under different numbers of known cycles.

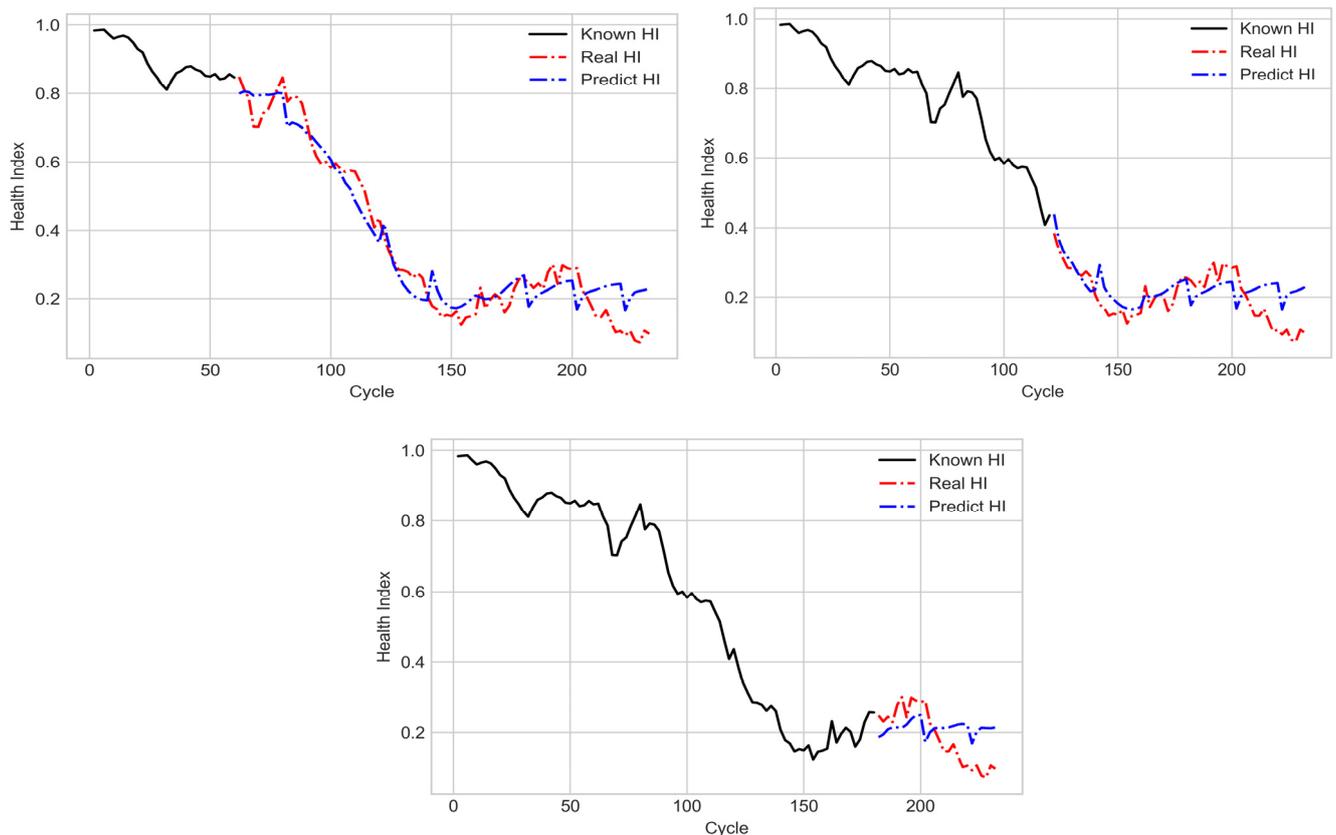|	30 Known Cycles	60 Known Cycles	90 Known Cycles
Training dataset	0.006889	0.005170	0.009435
Testing dataset	0.005132	0.004260	0.007701

Figure 21. Predicted degradation trajectory for the testing dataset under different numbers of known cycles.

5. Conclusions

The conclusions of this study are as follows:

1. This study used a triaxial accelerometer to measure the vibration signals of dental handpieces under the free-running condition. The collected data were subjected to the FFT, after which the frequency spectrum of the rotation speed was plotted. Subsequently, the amplitudes of CFs and the main rotation speed frequency were identified in the frequency domain and compared to obtain two critical CFs related to the SOHs of the dental handpieces. These CFs were used as indices for predicting the RUL and assessing the degradation of the handpiece.
2. This study established an LSTM model with the many-to-one structure that can store and transfer memories and an LR prediction model that cannot store memories to predict the SOHs of the dental handpieces. These models exhibited favorable prediction results. However, the LSTM model had a lower mean square error than the LR model did, which indicates that the LSTM model could monitor the SOH for a long time. The aforementioned LSTM model can be used to predict the RULs of clinical dental handpieces with high accuracy.
3. An LSTM model with the many-to-many structure was used to predict the SOHs of unknown cycles. The obtained output was then used as the input for the next unknown cycle; thus, iterations were executed. Although the aforementioned LSTM model did not provide highly accurate results for the later cycles (with respect to the real HI), the model could favorably reveal the gradual degradation of the dental handpieces in the middle cycles.

Author Contributions: Conceptualization, Y.-C.H.; Investigation, Y.-C.H.; Project administration, Y.-C.H.; Software, Y.-H.C. All authors have read and agreed to the published version of the manuscript.

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

References

1. Saghlatoon, H.; Soleimani, M.; Moghimi, S.; Talebi, M. *An Experimental Investigation about the Heat Transfer Phenomenon in Human Teeth*; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2012; pp. 1598–1601.
2. Zakeri, V.; Arzanpour, S. Measurement and Analysis of Dental Handpiece Vibration for Real-Time Discrimination of Tooth Layers. *J. Sel. Areas Bioeng.* **2010**, *8*, 13–18.
3. Wei, M.; Dyson, J.E.; Darvell, B.W. Factors Affecting Dental Air-Turbine Handpiece Bearing Failure. *Oper. Dent.* **2012**, *37*, E1–E12. [CrossRef]
4. Wei, M.; Dyson, J.E.; Darvell, B.W. Failure analysis of the ball bearings of dental air turbine handpieces. *Aust. Dent. J.* **2013**, *58*, 514–521. [CrossRef] [PubMed]
5. Huang, Y.-C.; Liu, C.-C.; Chuo, P.-C. Prognostic Diagnosis of the Health Status of an Air-Turbine Dental Handpiece Rotor by Using Sound and Vibration Signals. *J. Vibroeng.* **2016**, *18*, 1514–1524.
6. Xu, L.; Pennacchi, P.; Chatterton, S. A new method for the estimation of bearing health state and remaining useful life based on the moving average cross-correlation of power spectral density. *Mech. Syst. Signal Process.* **2020**, *139*, 106617. [CrossRef]
7. Ren, L.; Sun, Y.; Cui, J.; Zhang, L. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. *J. Manuf. Syst.* **2018**, *48*, 71–77. [CrossRef]
8. Beganovic, N.; Söffker, D. Remaining lifetime modeling using State-of-Health estimation. *Mech. Syst. Signal Process.* **2017**, *92*, 107–123. [CrossRef]
9. Yan, J.; Lee, J. Degradation Assessment and Fault Modes Classification Using Logistic Regression. *J. Manuf. Sci. Eng.* **2004**, *127*, 912–914. [CrossRef]
10. Zhang, S.; Zhai, B.; Guo, X.; Wang, K.; Peng, N.; Zhang, X. Synchronous estimation of state of health and remaining useful lifetime for lithium-ion battery using the incremental capacity and artificial neural networks. *J. Energy Storage* **2019**, *26*, 100951. [CrossRef]
11. Weng, C.; Sun, J.; Peng, H. A unified open-circuit-voltage model of lithium-ion batteries for state-of-charge estimation and state-of-health monitoring. *J. Power Sources* **2014**, *258*, 228–237. [CrossRef]
12. Zhao, D.; Wang, T.; Chu, F. Deep convolutional neural network based planet bearing fault classification. *Comput. Ind.* **2019**, *107*, 59–66. [CrossRef]
13. Zocca, V.; Spacagna, G.; Slater, D.; Roelants, P. *Python Deep Learning*; Packt Publishing: Birmingham, UK, 2017; Chapter 6.
14. Cen, Z.; Wang, J. Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer. *Energy* **2019**, *169*, 160–171. [CrossRef]
15. Zhang, J.; Wang, P.; Yan, R.; Gao, R.X. Long short-term memory for machine remaining life prediction. *J. Manuf. Syst.* **2018**, *48*, 78–86. [CrossRef]
16. Elsheikh, A.; Yacout, S.; Ouali, M.-S. Bidirectional handshaking LSTM for remaining useful life prediction. *Neurocomputing* **2019**, *323*, 148–156. [CrossRef]
17. Heimes, F.O. Recurrent neural networks for remaining useful life estimation. In Proceedings of the 2008 International Conference on Prognostics and Health Management, Denver, CO, USA, 6–9 October 2008; pp. 1–6.
18. Kovvali, N.; Banavar, M.; Spanias, A. An Introduction to Kalman Filtering with MATLAB Examples. *Synth. Lect. Signal Process.* **2013**, *6*, 1–81. [CrossRef]