Research on Continuous Pipeline Life Prediction Method Based on Fully Connected Neural Network

Zhikui Zhang¹, Lina Wu²

¹School of Mechanical Engineering, Xi'an Shiyou University, Xi'an City, Shaanxi Province, 710065, China
²Linyi Dongxing Experimental School, Linyi City, Shandong Province, 276000, China

Abstract: Aiming at the low accuracy of traditional empirical formulas in predicting the fatigue life of continuous oil pipelines, a fully connected neural network is utilized to predict the low-week fatigue life of continuous oil pipelines. Considering the influence of internal pressure on the fatigue life of continuous oil pipeline during operation, a prediction method combining the fully connected neural network and gated recirculation unit is proposed, and the experiment proves that the FCNN-GRU neural network performs better in terms of prediction accuracy and stability compared with the BP neural network.

Keywords: Continuous oil pipeline; Life prediction; Fully connected neural network; Gated cyclic unit.

1. Introduction

Coiled tubing (CT) is made of low carbon alloy steel and is widely used in various fields such as well workover, drilling and logging in oil and gas fields [1]. In the actual working process, CT is prone to surface mechanical damage due to the friction and extrusion with the drum, injection head, well casing and other objects [2]. When the surface mechanical damage and fatigue load act on the tubing body at the same time, it leads to fracture failure of the continuous tubing in serious cases [3], and may even lead to significant economic losses. Therefore, timely and accurate prediction of the fatigue life of continuous oil pipelines is crucial to ensure the safe operation of continuous oil pipelines and oilfields.

Joanne Ishak et al.[4] showed that small defects occurring in oilfield environments can significantly reduce the fatigue life of continuous pipelines. A modified Neuber's rule was proposed using notched root strain results. Wan f et al.[5] An evaluation method combining in-line inspection technology and fatigue life can predict the remaining life more accurately. The actual defect size is further incorporated into the fatigue life evaluation calculation to eliminate the unreliability of the existing calculation model that does not consider the defect influence factor. Song Peng et al.[6] proposed a method to construct a fatigue life prediction model for continuous oil pipelines using LMBP artificial neural network (ANN). and the prediction accuracy is significantly better than the conventional modeling method. Yu Guijie [7] used SOFM and RBF to construct a hybrid network model to predict the life of continuous oil pipeline.. No scholars have conducted research on the fatigue life prediction of continuous oil pipeline based on fully connected neural network. In order to enrich the fatigue life prediction method of continuous oil pipeline, a network model based on FCNN-GRU network is established to predict the life of continuous oil pipeline.

2. FCNN-GRU Model Construction

The essence of FCNN computation is matrix operation, and the loss value is reduced by forward propagation computation and back propagation computation during the computation process, which has high accuracy [8].GRU can effectively solve the problem of gradient vanishing during the propagation process, and the structure of the network is simpler and has fewer training parameters compared to Long Short Term Memory (LSTM). The combination of FCNN and GRU can better handle flow pattern data and improve the accuracy of recognizing flow patterns.

2.1. FCNN model

FCNN is the most basic neural network structure, neurons are a nonlinear model with multiple inputs and single outputs. There is only one input and output layer in FCNN, which can contain multiple hidden layers as shown in the FCNN model diagram in Fig. 1.

\[
y_j^k = f \left( \sum_{i=1}^{n} w_{ij}^{k-1} y_i^{k-1} + b_j^k \right)
\]

Where \( w_{ij}^k \) is the weight occupied by the output of the ith node in layer k-1 when passed to the jth node in layer k, \( b_j^k \) is the bias corresponding to the jth node in layer k, and f is the activation function.

Activation function is a function that enhances the
nonlinearity of the network model, the introduction of nonlinear transformation as an activation function can map the input data to a higher dimensional feature space and improve the expressive ability of the network, the commonly used activation functions are mainly Sigmoid, Tanh, Relu, Leaky Relu, Softplus, and Mish, etc., and the activation function of each activation function is shown in Fig. 2.

![Figure 2. Plot of common activation functions](image)

Sigmoid activation function contains power operation and division, in the training process of computation is large, easy to appear gradient disappearance, can not complete the deep network training. Tanh activation function in the positive and negative semiaxis of the limit of the case tends to 1, the training is also easy to appear gradient disappearance. Relu activation function, although to a certain extent to improve the gradient disappearance and the gradient explosion of the problem, but in the backward propagation process, if the input is negative, the gradient will be completely zero, and the problem of neural death will occur. Leaky Relu activation function can solve the problem of neuron death, but due to its linearity, it can not be used for complex classification. Softplus, although it can be derived everywhere and can constrain the neuron output to be constant greater than 0, but it is not centered on 0 making the gradient update slower during training. Compared to other activation functions, the Mish activation function is overall a smooth function with non-monotonically smooth variations in the negative semiaxis allowing for better gradient flow, which ensures better generalization and effective optimization capabilities, with greater stability in deeper networks. The Mish activation function is:

\[
f(x) = x \tan h(\ln(1 + e^x))
\]

\[
\tan h(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}
\]

2.2. GRU model

GRU was first proposed by Cho et al to capture the dependencies between different time scales. GRU is similar in structure to LSTM, which can achieve almost the same prediction as the same, and is able to solve the problem of gradient explosion and disappearance that occurs in FCNN in practical applications. GRU contains only two gate structures, the update gate \( r_t \) and the reset gate \( z_t \). The structure of the GRU network unit is shown in Fig. 3. Compared with LSTM, the structure is simpler, requires fewer training parameters, and is faster to train.

![Figure 3. GRU network unit structure](image)

In the GRU network structure, \( x_t \) is the new input at the current moment and \( h_{t-1} \) is the hidden state at the previous moment, and the model equation is expressed as [26]:

\[
\begin{align*}
    z_t &= s \left( w^z x_t + w^b z_{t-1} + b_z \right) \\
    r_t &= s \left( w^r x_t + w^b r_{t-1} + b_r \right) \\
    \tilde{h}_t &= \tanh \left( w^h x_t + r_t \odot w^b h_{t-1} + b_h \right) \\
    h_t &= z_t \odot \tilde{h}_t + \left( 1 - z_t \right) \odot h_{t-1}
\end{align*}
\]

\( \tilde{h}_t \) is the candidate hidden layer state; \( h_t \) is the hidden layer state at time \( t \); \( \odot \) denotes the same-or operation; \( w^z, w^r, w^h \) is the weighting matrix for each door; \( b_z, b_r, b_h \) is the offset parameter of the respective gate; \( \sigma \) is a sigmoid activation function that can act as a gating signal.

2.3. Regularization techniques

Regularization technique refers to the introduction of additional information on the original learning model to solve the overfitting problem[9]. Commonly used regularization schemes mainly include L1, L2 and Dropout regularization techniques. The principle of Dropout regularization technique refers to randomly culling with a certain probability during each iteration of training, and using the network composed of the remaining neurons to train the sample data of this iteration, as shown in Figure 4. During the training process of fully connected neural networks, the Dropout mechanism is employed to significantly simplify the network structure and effectively prevent the neural network from overfitting, thus improving the training efficiency and prediction accuracy of the model. Therefore, the dependence of the prediction results on the hidden layer nodes is successfully reduced by applying the Dropout regularization technique to enhance the generalization ability of GRU.
3. Data set and Experimental Environment

3.1. Data set processing

The dataset used for the experiments is derived from the results of continuous pipeline fatigue tests in literature [10] and literature [7], and is summarized in a total of 20 sets of data, as shown in Table 1. numbered 1-15 is the training set, and 16-20 is the test set.

<table>
<thead>
<tr>
<th>sample number</th>
<th>Coiled tubing outer diameter /mm</th>
<th>Wall Thickness /mm</th>
<th>Bend radius /mm</th>
<th>Internal pressure /MPa</th>
<th>Test cycles frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.75</td>
<td>2.210</td>
<td>1219.0</td>
<td>1.72</td>
<td>522</td>
</tr>
<tr>
<td>2</td>
<td>31.75</td>
<td>2.210</td>
<td>1219.0</td>
<td>10.34</td>
<td>517</td>
</tr>
<tr>
<td>3</td>
<td>31.75</td>
<td>2.210</td>
<td>1219.0</td>
<td>27.58</td>
<td>212</td>
</tr>
<tr>
<td>4</td>
<td>31.75</td>
<td>2.210</td>
<td>1219.0</td>
<td>34.48</td>
<td>128</td>
</tr>
<tr>
<td>5</td>
<td>38.10</td>
<td>4.445</td>
<td>1828.8</td>
<td>31.03</td>
<td>325</td>
</tr>
<tr>
<td>6</td>
<td>38.10</td>
<td>4.445</td>
<td>1828.8</td>
<td>31.03</td>
<td>354</td>
</tr>
<tr>
<td>7</td>
<td>44.45</td>
<td>2.769</td>
<td>1828.8</td>
<td>27.58</td>
<td>93</td>
</tr>
<tr>
<td>8</td>
<td>44.45</td>
<td>3.962</td>
<td>1828.8</td>
<td>27.58</td>
<td>247</td>
</tr>
<tr>
<td>9</td>
<td>44.45</td>
<td>2.769</td>
<td>1219.0</td>
<td>1.72</td>
<td>362</td>
</tr>
<tr>
<td>10</td>
<td>44.45</td>
<td>2.769</td>
<td>1219.0</td>
<td>6.89</td>
<td>262</td>
</tr>
<tr>
<td>11</td>
<td>44.45</td>
<td>2.769</td>
<td>1219.0</td>
<td>13.79</td>
<td>216</td>
</tr>
<tr>
<td>12</td>
<td>44.45</td>
<td>2.769</td>
<td>1219.0</td>
<td>27.58</td>
<td>99</td>
</tr>
<tr>
<td>13</td>
<td>44.45</td>
<td>2.769</td>
<td>1219.0</td>
<td>34.48</td>
<td>49</td>
</tr>
<tr>
<td>14</td>
<td>44.45</td>
<td>3.962</td>
<td>1219.0</td>
<td>13.79</td>
<td>311</td>
</tr>
<tr>
<td>15</td>
<td>44.45</td>
<td>3.962</td>
<td>1219.0</td>
<td>27.58</td>
<td>235</td>
</tr>
<tr>
<td>16</td>
<td>60.32</td>
<td>3.404</td>
<td>1828.8</td>
<td>17.24</td>
<td>74</td>
</tr>
<tr>
<td>17</td>
<td>31.75</td>
<td>2.210</td>
<td>1219.0</td>
<td>20.68</td>
<td>333</td>
</tr>
<tr>
<td>18</td>
<td>38.10</td>
<td>3.962</td>
<td>1828.8</td>
<td>37.58</td>
<td>247</td>
</tr>
<tr>
<td>19</td>
<td>44.45</td>
<td>3.175</td>
<td>1828.8</td>
<td>31.03</td>
<td>325</td>
</tr>
<tr>
<td>20</td>
<td>44.45</td>
<td>2.769</td>
<td>1219.0</td>
<td>20.68</td>
<td>170</td>
</tr>
</tbody>
</table>

3.2. Experimental environment

<table>
<thead>
<tr>
<th>Name</th>
<th>Experimental Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>Windows11</td>
</tr>
<tr>
<td>Programming Language</td>
<td>Python</td>
</tr>
<tr>
<td>Deep Learning Framework</td>
<td>Pytorch</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel(R) Core(TM) i7-14700k</td>
</tr>
<tr>
<td>GPU</td>
<td>NVIDIA GeForce RTX 4060 12GB</td>
</tr>
<tr>
<td>CUDA</td>
<td>Cuda 12.1</td>
</tr>
<tr>
<td>Flat-roofed Building</td>
<td>Pycharm 2022.3</td>
</tr>
</tbody>
</table>

3.3. Network training

The hyperparameter settings for this experiment have an initial learning rate of 0.01, momentum of 0.946, training number of 200, and a weight decay rate of 0.0005 using the SGD optimizer.

4. Experimental Results and Analysis

The evaluation metrics used in this experiment are Precision P (Precision), Recall R (Recall), Average Precision AP (Average Precision), Mean Average Precision mAP (meanAP,mAP), and the related evaluation metrics are:

\[
P \text{ (Precision)} = \frac{TP}{TP + FP} \tag{5}
\]
\[
R \text{ (Recall)} = \frac{TP}{TP+FN} \quad (6)
\]
\[
AP = \sum_{i=1}^{n} P(i) \Delta R(i) = \int_{0}^{1} P(R) dR \quad (7)
\]
\[
mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \quad (8)
\]

With the continuous training of the model, the localization loss (box_loss) and classification loss (cls_loss) of the training set continue to decrease, and the confidence loss (obj_loss) of the training set stabilizes at around 0.02 after 100 rounds of training. The localization loss and confidence loss of the validation set stabilized at around 0.0025 and 0.003 respectively after 150 rounds of training, and the classification loss stabilized at around 0.001. The improved model has a detection accuracy close to 0.98 after 150 rounds of training.

![Figure 5. FCNN-GRU model effect diagram](image)

### 4.1. Two network prediction results

In order to test the comprehensive performance of the FCNN-GRU algorithm, a BP neural network was chosen to train and test the FCNN-GRU fractal under the same continuous pipeline fatigue dataset, training set, and experimental strategy, and the network test results were obtained as shown in Fig. 6.

![Figure 6. Neural network prediction results for samples #8 and #12](image)

As can be seen from Figure 6, the two neural networks do 15 experiments under the conditions of the same provision of training samples and experimental strategies, the test results of FCNN-GRU's network are stable, while the test results of the BP neural network fluctuate up and down around the experimental values, which indicates that the overall stability of the FCNN-GRU network is better than that of the BP neural network, and thus the FCNN-GRU neural network has good accuracy and generalization. FCNN neural networks have the advantages of weight sharing and local awareness, but FCNN has some disadvantages in dealing with time series data, while gated recurrent unit (GRU) can deal with the problem of insufficient dependence on long distance of time series.
5. Concluding

A neural network model combining a fully connected neural network (FCNN) and a gated recurrent unit (GRU) is applied to predict the fatigue life of a continuous oil pipe. Among them, the GRU, as a variant of the RNN, has a memory function and incorporates temporal features, so that the FCNN-GRU model is able to capture the time nodes at which the degradation characteristics of the continuous pipeline life appear, and to improve the prediction accuracy of the remaining service life of the continuous pipeline.

References