MFD: Multi-object Frequency Feature Recognition and State Detection Based on RFID-single Tag

BIAOKAI ZHU, Shanxi Police College, P.R.China
ZEJIAO YANG, Shanxi Police College, P.R.China
YUPENG JIA, Shanxi Police College, P.R.China
SHENGXIN CHEN, Shanxi Police College, P.R.China
JIE SONG*, Sichuan Police College, P.R.China
SANMAN LIU†, Shanxi Police College, P.R.China
PING LI, Anhui University, P.R.China
FENG LI, Taiyuan University of Technology, P.R.China
DENG-AO LI, Taiyuan University of Technology, P.R.China

Vibration is a normal reaction that occurs during the operation of machinery and is very common in industrial systems. How to turn fine-grained vibration perception into visualization, and further predict mechanical failures and reduce property losses based on visual vibration information, which has aroused our thinking. In this paper, the phase information generated by the tag is processed and analyzed, and MFD is proposed, a real-time vibration monitoring and fault-sensing discrimination system. MFD extracts phase information from the original RF signal and converts it into a markov transition map by introducing White Gaussian Noise and a low-pass filter for denoising. To accurately predict the failure of machinery, a deep and machine learning model is introduced to calculate the accuracy of failure analysis, realizing real-time monitoring and fault judgment. The test results show that the average recognition accuracy of vibration can reach 96.07%, and the average recognition accuracy of forward rotation, reverse rotation, oil spill, and screw loosening of motor equipment during long-term operation can reach 98.53%, 99.44%, 97.87%, and 99.91%, respectively, with high robustness.

CCS Concepts: • **Computer systems organization** → **Downlink transmission**; *Multicast data distribution*; • **Software and its engineering** → Software configuration and firmware update.

Additional Key Words and Phrases: IoT, RFID, Passive Sensing, Contactless State Detection;

*This is the corresponding author
†This is the corresponding author

Authors' addresses: Biaokai Zhu, Shanxi Police College, Taiyuan, P.R.China, hongtaozhuty@gmail.com; Zejiao Yang, Shanxi Police College, Taiyuan, P.R.China, yangzejiao2001@163.com; Yupeng Jia, Shanxi Police College, Taiyuan, P.R.China, 3175264800@qq.com; Shengxin Chen, Shanxi Police College, Taiyuan, P.R.China, 940245785@qq.com; Jie Song, Sichuan Police College, No. 186, Longtouguan Road, Jiangyang District, Luzhou, Sichuan, 646000, P.R.China, scjcxysj@163.com; sanman Liu, Shanxi Police College, P.R.China, lsm601719@126.com; Ping Li, Anhui University, P.R.China, 20145@ahu.edu.cn; Feng Li, Taiyuan University of Technology, P.R.China, lizhuzhu09@163.com; Deng-ao Li, Taiyuan University of Technology, P.R.China, lidengao@tyut.edu.cn.

2577-6207/2023/8-ART $15.00
https://doi.org/10.1145/3615665

1 INTRODUCTION

With the rapid development of the Internet industry, the world's requirements for essential support and technical breakthrough are also getting higher and higher. The premise of completing these requirements is to ensure high integration between information technology and industry system. In intelligent industrial systems, rotation has been widely used in various fields, such as smart manufacturing, medicine, home appliances, transportation, etc. Rotating refers to the circular motion of a device around a point or axis. Mechanical wheel[21] is the core part of industrial production. For example, from small to large turntables[1], it can apply to all significant industrial transportation chains. Mechanical rotation produces vibration, and automated vibration signals can reflect the state of automatic operation, such as (Fig.1), the vibration signals existing in the fan[6], motor[20], gear[23] , and unscrewed aerial vehicle (UAV)[28] can reflect the operation of the objects and their running state feedback. The era of speed detection of single and straightforward vibration equipment has passed, and multiple mixing equipment detections have become a routine deployment in the industry field. Breakthrough the limitations of a traditional single device's speed detection and distinguishing the signal and speed of the mixing equipment has become an important step. The fault feature extraction for mechanical rotating is one of the critical steps to determine whether a mechanical failure. How to ensure its safe and stable operation has great reality significance on the industry of the whole system.

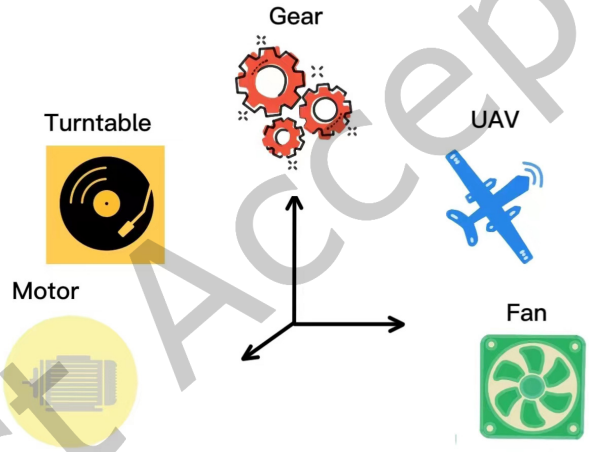


Fig. 1. Serveral potential vibration

Traditional methods of equipment detection mainly include perception methods based on special sensors, such as acceleration, piezoelectric, electromagnetic, etc. Most of these methods refer to complex, time-consuming, and inexact invasive deployment. In addition, the installation of sensors may cause downtime and property losses. Video-based perception methods, such as high-speed cameras, laser tachometers, etc. They have specific requirements for light, and it is difficult to perceive the movement of objects in the case of insufficient light or occlusion.These traditional sensing methods can effectively convert the rotation information into electronic information, laying the foundation for accurately predicting the rotation speed in the following stage.

It can be seen from the above that in the research progress of the industrial field, the proposed sensing methods all have one commonality: binding deployment on rotating devices, and there is no lack of low-cost sensing methods based on RFID, such as TAGBEAT, TAGTWINS, RED, TAGSOUND and so on. Because of this, it also brings a series of problems. Firstly, the binding deployment of tags may cause equipment to fall off in high-speed

mechanical rotation, low safety performance, and expensive costs due to the waste of workforce and the property damage caused by downtime. Secondly, the deployment of precision instrument tags has become a problem. The binding deployment has a significant impact on the accuracy of the instrument. Finally, due to the long-term high-speed rotation of the equipment, the tag may be loose during the rotation process. Still, the kit does not produce a warning and manually fix the tag, leading to the results being concluded under inconsistent conditions of the experimental scene, which significantly reduces the versatility and accuracy of the detection speed.

To achieve the real-time detection of motor fault based on a single tag, we propose an innovative noncontact perception detection system, which uses the simulated mechanical operation state composed of an antenna, reader, tag, and motor, and uses the phase information generated by the tag to realize the perception recognition and analysis of motor vibration frequency and fault. We currently divide the blame into four categories: regular operation, motor inversion, oil infiltration, and screw loosening. We also pose the following challenges:

- **High-frequency signal recovery without distortion.** The traditional vibration signal extraction method has a low sampling rate, which cannot meet the needs of sensing the running state of the machinery in the actual production and life. The non-contact sensing method used in the system weakens the interference of external factors and realizes real-time sensing, but how to accurately sense the vibration frequency and transmit the collected sample signal without distortion has become a challenge.
- **Separation of multi-object vibration signals.** In the actual application scenario, due to the interference of the surrounding environment, equipment and personnel, the traditional signal detection method has low robustness. It cannot accurately identify the signal frequency corresponding to the vibrating object. Since the basic frequency recognition is oriented to a single object, in the face of mixed signals generated by multiple objects vibrating at the same time, how to accurately match the signal and its corresponding high-frequency signal and reduce the harm of the harmonic caused by the resonance of the object has become a challenge.
- **The accuracy of identifying the characteristics of different fault states.** In the long-term operation of machinery, the motor may occur screw loosening and other faults. However, these state characteristics are fine-grained, but RFID radio frequency technology identifies coarse-grained information and cannot accurately detect and identify the fault. Therefore, how to extract the phase feature of the vibration signal with high precision and provide favorable conditions for the subsequent neural network fault classification and identification has become a challenge.

To solve the first challenge, we designed a contactless intelligent sensing system based on a single tag. The non-contact deployment of the tag makes the function of perception and recognition without destroying the mechanical operation. The traditional signal acquisition method follows the Nyquist Sampling Theorem, which requires the sampling rate to be twice the highest frequency. Still, the sampling theorem cannot guarantee the signal without distortion. We use the improved compressive sensing technology to effectively collect the high-frequency signal, solve the problem of limited distortion of the sampling theorem and successfully recover the band-limited signal.

To solve the second challenge, when multiple objects work simultaneously, we use Fourier Series Decomposition to decide whether harmful harmonics are generated and whether these harmonics belong to what harmonics. When multiple object signal sources appear, signal superposition will happen, which is easy to produce a coupling phenomenon and not easy to match the signal. Wigner-Ville transform is used to process the non-stationary signal, and the object's vibration is accurately recognized.

To solve the third problem, we adopt the Markov Transition Field to amplify the one-dimensional time-series signal features collected by the reader and use the FIR filter based on the Fast Fourier Transformation. By reducing the sampling rate to 1/4 of the original frequency when sampling the signals within the specified range, separation

and purification are achieved, and the coarse-grained features are fine-grained to achieve the effect of fault recognition and classification.

To solve the above challenges, we design a non-contact vibration sensing and fault monitoring system based on a single tag, taking the vibration equipment motor as an example. The specific idea is as follows: Our experimental deployment is straightforward. An antenna and RFID tag are put around the vibration device, and the signal sensed by the antenna is transmitted to the reader. The reader reads data to recognize the vibration signal's perception recognition. In detail, our method removes the DC component of the collected tag phase information and process the data by adding Gaussian Noise and a low-pass filter to realize the frequency identification of the device vibration. Firstly, the phase information after data processing is used to recover the high-frequency signal without distortion using the improved basis pursuit denoising algorithm. Secondly, the Fourier Series Decomposition was used to determine whether there was a harmonic generation and confirm the number of objects in the multi-signal model. Then the signal was matched with the objects one by one. Finally, the phase information of the signal was converted to the signal image, which was trained in the improved convolutional neural network SVGG to obtain the accuracy of fault detection. To sum up, the contributions of this paper are as follows:

- Lossless recovery of high-frequency signals. We use an improved basis pursuit denoising algorithm to compress and reconstruct the high-frequency signal so that the noise is reduced in the process of compression to ensure the accurate recovery of the high-frequency signal.
- We judge the existence of harmonics by Fourier Decomposition, use the Wigner-Ville distribution to suppress the cross interference generated by the multi-signal components, analyze the time-frequency domain information of objects, and realize the signal matching of the multi-signal system model with high precision.
- We design a signal matching and fault real-time detection system based on a single tag, which uses commercial readers, antennas, a single tag and multiple motors to simplify the experimental deployment. The improved convolutional neural network is used to train and learn four types of faults, and the accuracy rate reaches more than 95%, fully proving the system's robustness.

2 RELATED WORK

Rotation sensing detection is a multidisciplinary synthesis. The rotation state of industrial production equipment critically influences its efficient and safe production. The separation and detection of multi-target rotation signals is the premise of accurate fault diagnosis and breaking the traditional one-dimensional velocity detection. Multiple overlapping particular combinations are difficult to detect in time for rotating machinery in power generation, mining, aviation and other industrial systems. Although early damage detection is crucial, the rotation information of multiple mixed signals is often more challenging to separate and purify than the single information and is greatly affected by external factors.

2.1 Diagnostic method based on dedicated sensor

In traditional sensing systems, researchers often install special sensors on the equipment to sense the operating state parameters of objects, such as vibration information[22], acoustic features[8], audio features[11], thermal features[2], speed[29], etc. These research methods are not universal, and there are certain constraints in the actual scene. Z.qu used a highly sensitive optical accelerometer to measure the in-plane motion of mass by taking advantage of its insensitivity to electromagnetic interference[14].

Because an accelerometer can only sense high-frequency signal, low-frequency weak signal is not sensitive. K. Yamashita et al. used the piezoelectric sensor to process the low-frequency microwave array and determine

the 3D position of the object by mechanical tuning[24]. Although these efforts effectively sense the target state, deploying specialized sensors is complex, the application scope is limited, and some measures are intrusive.

2.2 Diagnostic methods based on visual AIDS

In addition to deploying specialized sensors, visual AIDS is often used in the perceptual domain. For example, Julien N. P. Martel used ordinary low-frame cameras to optimize functions end-to-end and recover information from scenes to perform high-speed compression imaging[12]. Y. Xu et al. used static background or dynamic cyclic motion in space to calculate image interpolation, improve image resolution, and detect motion state through spatial consistency[17]. Y. Aytar estimated the time of repeated operations in the video and, after in-depth model training, calculated the period and number of repeats in the video to capture the moment and the number of failures in the vibration process[4]. Zhang H divided the video data into periods and input it continuously to predict the motion state of the object the next time to judge whether there will be abnormalities[27]. However, the implementation of this approach has specific requirements. Once the light is insufficient or the measured object is occluded, the perception cannot proceed.

2.3 Non-contact diagnostic method

With the popularity of passive sensing technology, non-contact sensing has attracted extensive attention due to its simple deployment, wide application range, and high accuracy. For example, Xie B used TAGSMM to measure spokes with sub-millimeter resolution and proposed a phase change amplification caused by coupling vibration to detect motor screw loss effectively[18]. Yang P used a single tag in RF-EAR to accurately monitor the vibration of multiple devices and distinguish devices for identity authentication[25]. D L uses frequency-modulated continuous waves to analyze vibration parameters and achieve subhertz and micron amplitude accuracy[3]. Gil H uses ultrasonic tactile perception to realize object recognition, positioning, and state perception[5]. Xu C proposed WaveEar[19], an end-to-end anti-noise sensing system that uses millimeter-wave signals to induce vibration in complex environments. The system extends the range of vibration identification and motion condition monitoring.

2.4 Diagnostic methods based on signal analysis

Fault information in vibration signals is often manifested in frequency, and the fault information can be determined by identifying and analyzing frequency. Based on this, scholars analyze and study vibration signals. Weiqiang Yao[26] et al. designed a signal reconstruction algorithm to reconstruct detected signals by reconstructing the coefficients of signs of various scales to highlight fault features. Feng Tian[16] et al. used integrated empirical mode decomposition to decompose the original signal and divide it into stationary and non-stationary parts for feature extraction to establish fault diagnosis models. Li, Yuxing, et al. developed an LZC (DELZC) based on dispersion entropy to effectively capture the dynamic changes of time series and has advantages in characterizing signal complexity[10]. Kafeel et al. collected 3D vibration signal data sets[7]. He concluded through practice that the fault classification performance based on the Gaussian kernel support vector machine was the best by integrating the time feature and spectral feature.

2.5 Diagnosis method based on time-frequency domain analysis

In many practical applications, the frequency spectrum of a vibration signal often changes with time, so it is necessary to find out how the signal's energy is distributed in the time-frequency plane. R. Li[9] et al. proposed a fault feature gene extraction algorithm based on time-frequency domain statistics to establish a gene bank for fault detection. W. Mengjiao[13] et al. proposed establishing a fault diagnosis model based on time-domain and frequency-domain and time-frequency features after integrated empirical mode decomposition (EEMD),

which has a high classification accuracy. X. Song[15] et al. proposed a distribution network grounding fault line selection method based on a time-frequency domain energy matrix with high adaptability and reliability.

3 PRELIMINARY

3.1 Tag Deployment

In the RFID contactless system we designed, due to the fragile signal characteristics read by the reader, this paper studied the advantages and disadvantages of single-tag data acquisition, double-tag data acquisition and multi-tag data acquisition before the experiment. The signal data collected by a single tag has fragile signal characteristics,but the signal dada under a single tag is easy to assemble. It requires fewer environmental factors to be considered and is suitable for the condition that the signal is challenging to distinguish. When collecting double tags, the first thing to consider is the placement of the double tags. Due to the need to consider confrontation, diagonal, superposition and different spatial arrangements, the experimental operation is relatively complicated. In the signal processing process, the coupling effect between tags and the interference between objects should be observed to confirm the accuracy of the tag signal. When the multi-tag collection is used to process the collected signals, select several tag signals first. In essence, it is similar to a double tag, but the double tag is more representative, a special multi-tag collection method. In the preliminary exploration process, the selected tags are made more representative to maximize the benefits of single-tag deployment. The tagging characteristic introduces the game theory antagonism into the tag selection process. Still, the use of antagonism is very complicated in this process.

Although the signal obtained can reach the best in theory, it is hindered by interference factors and the application of game theory in reality. This paper selected the label between the motor and the antenna, and through the design of vibration object, antenna, reader, and tag layout, find out the best signal generation of experimental equipment layout. In this experiment, we deployed the experimental device integrating the motor, antenna, tag and reader to observe that the noise interference of the surrounding environment significantly impacts the signal generated by the vibration of the motor. Under different RF signal propagation paths, the same signal will fluctuate from transmitter to receiver. Generally, signal fluctuation refers to the signal difference produced by various environmental disturbances. In radio frequency identification, the propagation modes of radio frequency signals include reflection, scattering, refraction, diffraction, free-space path loss and so on. Finally, the signal energy received by the reader and antenna is enhanced and attenuated. To eliminate interference without affecting mechanical motion and reduce the influence of the multipath effect, the plugging of different materials is added,and the type and orientation of the tag are changed. The influence of the multipath effect is analyzed from different angles, and the signal waves of multiple motors are superimposed to make them better perceive the object characteristics.

3.2 Selection of signal features

For the generality of the system, we use a commercial reader to extract tag signals. Commercial readers do not provide the underlying data interface, especially the physical layer data that the tag communicates with the reader cannot be obtained. The reflected tag signal received by the commercial reader contains only the received signal strength indication, Phase Angle value and Doppler translation data. The value of these reflected signals changes with the location of the tag, the speed of movement and the surrounding environment.

$$P_{R\_t} = P_{T\_r} G_r G_t \alpha \left(\frac{\lambda}{4\pi d}\right)^2 \tag{1}$$

The reader sends a continuous electromagnetic wave through the antenna to the tag, and the tag chooses whether to backscatter the received electromagnetic wave back to the reader. According to the radio electromagnetic wave signal attenuation formula, the electromagnetic wave energy received by the tag is:

$$P_{R\_r} = \beta P_{T\_r} G_r^2 G_t^2 \alpha^2 (\frac{\lambda}{4\pi d})^4 \tag{2}$$

On type, β for the tag to receive reader signals of energy utilization, P_{T_T} on behalf of the reader sends out energy, G_r on behalf of the reader antenna signal gain, G_t on behalf of the tag signal gain, α on behalf of the transmission channel attenuation coefficient between reader and tag, λ as the radio wave wavelength, d as the communication distance between tag and reader. As can be seen from the above equation, the signal energy of the reader receiving the tag and returning it to the reader will decay with high intensity to the fourth power as the communication distance increases. In order to improve the ability of the reader to receive the weak electromagnetic wave signal reflected from the tag, we use a circularly polarized antenna. The large size and high gain of the antenna can compensate for the high intensity signal attenuation due to the distance.

3.2.1 Doppler shift. Doppler shift refers to the frequency difference between the transmitted signal and the received signal caused by Doppler effect. This phenomenon exists between all kinds of waves. Different colors of light waves represent different frequencies, which is the best confirmation of the Doppler shift phenomenon. The Doppler effect shows that the wavelength of radiation from an object change under the relative motion of the source and the observer. Commercial readers can directly measure the signal strength and phase value of the tag. The reader obtains the Doppler offset by calculating the phase difference at different times. Assume that the moving rate of the tag relative to the reader antenna is v , the Angle between the velocity vector and the length vector is α , and the signal frequency change of the reader antenna relative to the tag is $\frac{v}{\lambda}\cos(\alpha)$. Similarly, compared with the reader, the change of tag signal frequency is $\frac{v}{\lambda}\cos(\alpha)$. Therefore, the tag signal frequency received by the reader antenna relative to the transmitted signal frequency can be summarized as follows:

$$f_D = \frac{2v}{\lambda}\cos(\alpha) \tag{3}$$

3.2.2 RSSI values. The Received Signal Strength indicator (RSSI) is a measure of the energy of the tag signal received by the reader. The relationship between RSSI and signal energy can be expressed as follows:

$$RSSI = 10\log_{10}(\frac{P}{1mW}) \tag{4}$$

For RF sign, we use a commercial card reader with a channel attenuation coefficient of , which is usually 4 in the system. In the actual situation, the value of RSSI changes, from strong to weak values corresponding to 0dBm and -115dbm respectively. For the accuracy of the experiment, we usually set the distance between the card reader and the tag within 1m. When the value of is set to 1, substitute Eq.4 into Eq.2, and it can be simplified as follows:

$$RSSI = RSSI_{d=1} - 10\gamma\log_{10}(d) \tag{5}$$

3.2.3 Phase values. Phase is a particular point on a waveform in periodic motion. Phase information refers to the phase difference between the transmitted signal and the received signal. Two electromagnetic waves of the same frequency and phase can be superimposed. Similarly, two electromagnetic waves of the same frequency and reverse, that is, electromagnetic waves with a phase difference of π are superimposed to cancel out. The phase information is constantly transformed in the range $(0, 2\pi)$, which can be expressed as:

$$\theta = \left(\frac{d}{\lambda}\right) \bmod (2\pi) \tag{6}$$

When the relative distance between the reader antenna and the tag and the electromagnetic wave scattering distance change, the phase information of the tag signal will also change. Phase information is suitable for more fine-grained perception. If the communication distance between the tag and the reader antenna is d, λ represents the transmission wavelength, θ_{device} represents the system noise generated by the tag and reader hardware. Then the phase information of the tag received by the reader can be expressed as:

$$\theta = \left(2\pi \frac{2d}{\lambda} + \theta_{device}\right) \bmod (2\pi) \tag{7}$$

To achieve the innovation and versatility of the system, we built the system and collected signals through commercial card readers. Commercial card readers can receive signal indicators, including RSSI, Doppler, and Phase. The phase signal is selected from the three tag characteristic signals for analysis by observation. A phase is a wave that has peaks and troughs at specific locations. It is a measure used to analyze changes from one state to another, showing periodic changes. User actions can be analyzed by phase wave analysis. Therefore, the phase signal acquisition is carried out in the case of single tag and single antenna deployment.

4 CORE DESIGN

4.1 Data Collection

Sampling is the process of converting continuous signals into numerical sequences. Reasonable discrete signal sampling can retain the relevant information of the original signal. The Nyquist Sampling Theorem followed by commercial readers proposes that: when the sampling frequency is greater than equal to twice the bandwidth of the original signal, the original signal can be restored accurately through the collected signal. On the contrary, the high-frequency signal in the signal will be distorted, and the phenomenon of cross and hybrid will occur, further increasing the difficulty of recovery. The vibration intensity is tremendous in a natural environment, while the sampling rate of the ordinary reader is about $40Hz$, and the frequency of the collected signal, according to the sampling theorem, can only reach $20Hz$ or even lower, which does not meet the needs of the industrial field.

Compressed Sensing (CS), as an emerging sampling theory, can break the limits of traditional sampling theorems and accurately recover sparse signals that are compressed or known variation domains through nonlinear reconstruction algorithms. Among them, compressed sensing must meet the prerequisite conditions of sparsity and irrelevance. First of all, the signal should have the characteristic of sparse expression in a particular change domain; that is, in a specific change domain, it has fewer non-zero numbers. If the rotating signal $S(t)$ has periodic characteristics in the time domain, then the signal will be sparse in the frequency domain after FFT, namely.

$$\mathrm{S} = \psi \mathrm{s} \tag{8}$$

$$y1 = h(x) + F(x_l, W_l) \tag{9}$$

Where, S is the sparsity coefficient of the signal in the frequency domain, and matrix is the Fourier basis matrix. Therefore, sparse expressions of correlated signals in frequency domain are obtained. Secondly, there is no correlation between the observation matrix and the sparse representation basis requirements. Based on the two premises, the CS theory is mainly to select n sampled signals x with S sparsity on appropriate sparse basis ψ, and then obtain accurate reconstruction by linear projection $y(i) = \left\langle X, \phi_i^T \right\rangle, i \in \{1, 2, \cdots, m\}$ of the signals on m $(S \leq m \leq n)$ another incoherent basis $\phi = \left(\phi_1^T, \phi_2^T \cdots \phi_m^T\right)$

$$y = \phi x = \phi \psi^T S \tag{10}$$

Where, ϕ is the measurement basis and the n-dimensional matrix formed by it is the measurement matrix.$\phi\varphi^T$ is the sensor system matrix, and y is the measurement vector. For the method of recovering x from S , this paper mainly uses convex optimization algorithm to recover high-frequency signals.

Convex optimization algorithm is realized based on minimum norm model, including Basis Pursuit algorithm and Bregman algorithm. Greedy algorithm is based on Iterative computing, including Orthogonal Matching Pursuit (OMP) and Iterative Hard range algorithm IHT (Iterative Hard Thresholding). These methods have certain strict requirements on the perception matrix A. That is, the perception matrix A must satisfy the finite equidistant property (RIP). To be precise, let support $\wedge \subset \{1, \cdots, d\}$, a submatrix $\psi_\wedge$ of matrix ψ consisting of indexed columns, have a minimum local RIP constant $\delta_\wedge = \delta_\wedge(\psi)$ and satisfy the inequality.

$$(1 - \delta_\wedge)||\mathrm{x}||_2^2 \leq ||\psi_\wedge x||_2^2 \leq (1 + \delta_\wedge)||x||_2^2 \tag{11}$$

Where, $\delta_\wedge$ is called the $\wedge$ order equidistant constraint constant, if the perception matrix A satisfies the RIP of a certain order and constant, the optimal solution of the L-0 norm and the L-2 norm will be exactly the same.

The global RIP constant is:

$$\delta_s = \delta_s(\psi) := \sup_{|\wedge|=s} \delta_\wedge(\psi), S \in N \tag{12}$$

Based on the above theories and formulas, this system proposes that to obtain more accurate original signals, an iterative soft threshold algorithm (IST) should be used to solve the coefficient solutions of linear equations and the sparse solutions of nonlinear constraint problems, so as to recover the three-dimensional vibration signals. The IST algorithm does not keep the signal above the threshold but shrinks it appropriately. In short, when the soft threshold algorithm is used for threshold operation, threshold iterative operation should be carried out on a certain point until the value of this point tends to be stable. Threshold shrinkage operations are as follows:

$$S_\lambda = \begin{cases} (|xi| - w * \lambda) * sign(xi) & |xi| > w * \lambda \\ 0 & |\mathrm{xi}| < \mathrm{w}*\lambda \end{cases} \tag{13}$$

Where, x represents signal data in sparse , λ is the threshold that needs to be applied, which changes as the data changes. On this basis, the base tracking denoising algorithm (BPDN) is introduced to denoise the signal. This method proposes to use l1 norm instead of l0 norm to solve optimization problems so that linear programming method can be used to solve. Compared with the traditional denoising methods, the base tracking denoising method can improve the signal-to-noise ratio to a higher degree, and at the same time, it can effectively preserve the important features of the signal. The specific steps of base tracking denoising are as follows:

Basis Pursuit (BP), which proposes to use l_1 norm instead of l_0 norm to solve the optimization problem, so that linear programming method can be used to solve the problem. The steps of base tracking denoising are as follows:

First, a smooth approximation of the norm is given as follows:

$$\|z\| \approx \sum_{i=1}^{N} \left(|(z)_i|^2 + \varepsilon\right)^{1/2} \tag{14}$$

Where, N is the length of z the vector, $\varepsilon$is a very small constant,

$$J(\alpha) = \frac{1}{2}(y - \emptyset\alpha) + \lambda \sum_{i=1}^{N} (|\alpha_i|^2 + \varepsilon)^{12} \tag{15}$$

$$\mathrm{D}(p||q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = E_{p(x)} \log \frac{P(x)}{q(x)} \tag{16}$$

Solve the above equation iteratively, and the iterative formula is as follows:

$$\overset{\wedge}{\alpha}^{(n+1)} = \beta\left(\emptyset^H\emptyset + \lambda A_1\left(\overset{\wedge}{\alpha}^{|n|}\right)\right) - \emptyset^H y + (1-\beta)\,\overset{\wedge}{\alpha}^{|n|} \tag{17}$$

Where, β is the iteration step, and the initial value of iteration should be $\alpha(0) = \varphi HY$. The termination condition of the iteration is given by control.

$$\left\|\overset{\wedge}{\alpha}^{(n+1)} - \alpha^{(n)}\right\|_2^2 \Big/ \left\|\alpha^{(n)}\right\|^2 < \delta_{CG} \tag{18}$$

ALGORITHM 1: algorithm caption

Input:

x:data
A, At:function handles for A and its conj transpose
p:Parseval constant
λ:regularization parameter
μ:ADMM parameter
Nit:Number of iterations

Output:

x:solution to BPD problem

```
if n arg out < 2 then
    ComputeCost = False
else
    ComputeCost = true
    cost = zeros(1, Nit)
end
x = At(y);
d = zeros(size(x));
for i = 1:Nit do
     u = soft( x + d ,0.5 * λμ ) - d
    d = 1 / ( λ + p ) * At( y - A(u) )
    x = d + u;
    if n arg out > 2 then
        residual = y - A(x)
        cost(i) = sum(abs(residual(:))^2) + sum(abs(λ * x(:)));
    end
end
```

In this way, solution of the optimization problem can be obtained. After obtaining the estimate α representing coefficient α , the reconstructed signal can be obtained from $s = \varphi\alpha$ .However, this method has an obvious denoising effect only on Gaussian white noise. In order to expand the application range, we improved it by modifying the noise constraint condition and replacing L1-norm with L2-norm. This method can realize high-frequency signal recovery quickly and stably, and it also has a good denoising effect.

As shown in Fig.2, the improved BP base tracking denoising algorithm can be used to reconstruct and restore the original signal. This method can better suppress sparse noise, expand the application range of the base tracking algorithm, and greatly improve the operation speed.

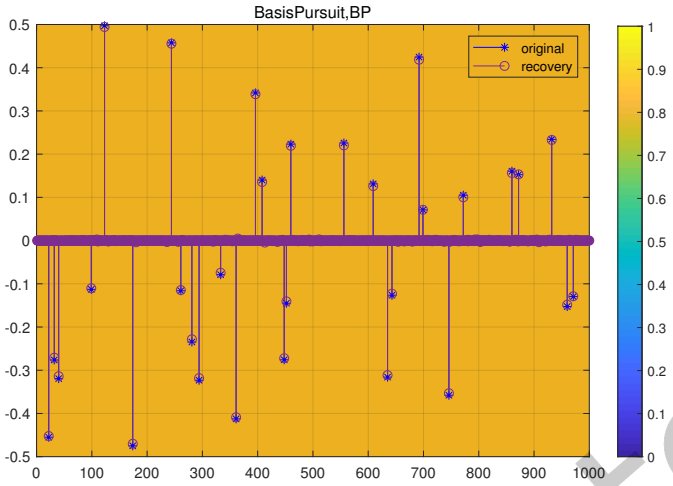


Fig. 2. BP high frequency signal recovery

4.2 Data Preprocessing

Signals in the process of collection and transmission, because of the influence of the external environment disturbance and the instrument itself, will have noise mixed in it. Noise is the impact signal detection and target recognition performance of an important factor, especially in some high-precision data analysis. Even a faint noise will have a significant impact on the analysis result. So the first thing we should consider is the possible purification of the data, that is the removal of excess noise by different methods, the analogy effect, and the selection of the best way. Rotor faults are common in rotating equipment. Due to the interference of external noise and the influence of the equipment itself, it is difficult to extract the fault characteristics of the fault signal accurately, so the rotor fault signal should be de-noised first. This paper compares the de-DC component and the sliding average de-noising method.

The DC component of the signal is the average of the signal, which is a constant independent of time. The DC component of the signal expressed by the formula is as follows:

$$f_{DC} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t)dt \tag{19}$$

If the original signal is periodic, the above equation can omit the process of taking the limit, and the limit of integration can take any period. The original signal was processed in addition to the DC component, and the processing results are shown in the Fig. 3 below:

In the moving average method, the signal information is given weight, the recent data is provided a large coefficient, and the long-term data is provided a small coefficient to eliminate the changing factors of the signal information. Dynamic test data are composed of deterministic components and random components, the former is the required measurement results or effective signals, and the latter is the test error or noise of random fluctuations. After discrete sampling, the dynamic test data can be written as follows:

By discretizing the deterministic and stochastic components of signal information, we can obtain:

$$y_i = f_i + e_i (j = 1, 2, \cdots, N) \tag{20}$$

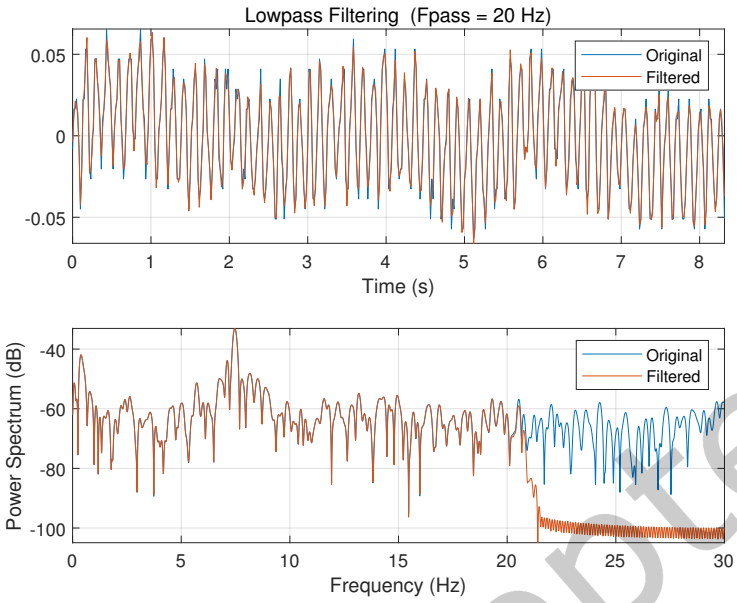


Fig. 3. Remove DC Component

In this paper, the original phase signal is processed by moving smooth filtering to reduce the influence of suddenly changing data. In order to achieve better actual denoising effect, the smoothing factor is set as 0.3 in the experimental setting, and the following Fig. 4 can be obtained.

By comprehensive comparison of the graph (DC component removal method) and graph (moving average method), it can be concluded that, compared with the moving average method, the DC component removal method has a flatter phase curve while retaining tag phase information, and the denoised curve has less jitter.

4.3 Identification of vibration frequency

The sampling rate of commercial readers is low, usually up to 40 Hz. According to the Nyquist Sampling Theorem and practical engineering results, we choose a Finite Impulse Response (FIR) filter implemented by FFT for filtering and vibration frequency identification. FIR low-pass filter can retain the phase frequency characteristics based on ensuring any amplitude-frequency features, and the unit sampling response is limited so that the system can maintain a relatively stable effect. The tag's phase information is easily disturbed by environmental noise, so a sound identification system is essential for vibration identification of equipment in complex environments.

FIR filter is a kind of approximation based on the frequency characteristics of ideal filter. It adopts the method of window function and sampling response of unit frequency to approach continuously. Assuming that the cut-off frequency of low-pass digital filter is $W_{cut-off}$ and the group delay is α, the unit impulse of filter should be expressed as follows:

$$h_d(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H_d(e^{jw})e^{jw}dw \tag{21}$$

Further derivation of the above formula shows that:

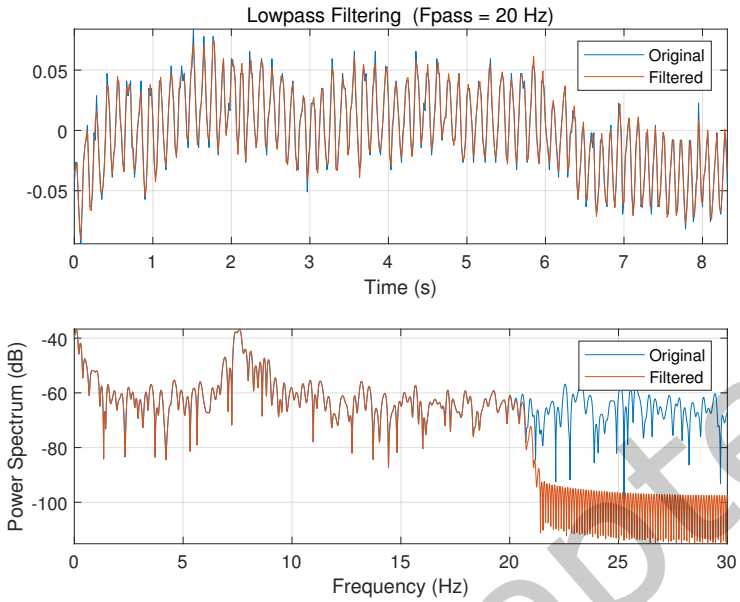


Fig. 4. Moving average denoising

$$h_d(n) = \frac{\sin(w_c(n-a))}{\pi(n-a)} \tag{22}$$

The unit impulse response of the ideal filter is infinite, but in practical application, the length of the low-pass filter is finite, so it is necessary to choose a better Hanning window function to adjust $h_d(n)$. According to the principle that half of the width of the main lobe is equal to the cutoff frequency, we estimate the length of the low-pass filter to realize high-precision vibration frequency perception.

4.4 Amplification of signal features

Model complexity and theoretical accuracy threshold of the multi-target vibration resolution method are greatly affected by data processing, so it is essential to process the original signal reasonably. On the one hand, data processing quality depends on signal processing methods to analyze signals from different devices generated and pick out key fingerprint features. On the other hand, convolutional neural networks are used to autonomously learn fingerprint data to realize the automatic extraction of high-quality components and classification of fingerprint features. This paper uses a single tag to simultaneously detect the vibration signals of multiple objects. Because the object vibration is very likely to occur in resonance phenomenon, producing harmonic, it will reduce the machine's working efficiency. Therefore, we need to detect whether or not the vibration of multiple objects has harmonics, early detection and avoidance.

Through Fourier Series variation, a weak harmonic response will be generated, and the amplitude of the harmonic component presented will also change with the increase of vibration frequency. Because the vibration signals of multiple objects are mixed signals, the signal variation trend of each object cannot be accurately obtained, and harmonics cannot be found in time and avoided perfectly. Harmonic is a hot topic in signal

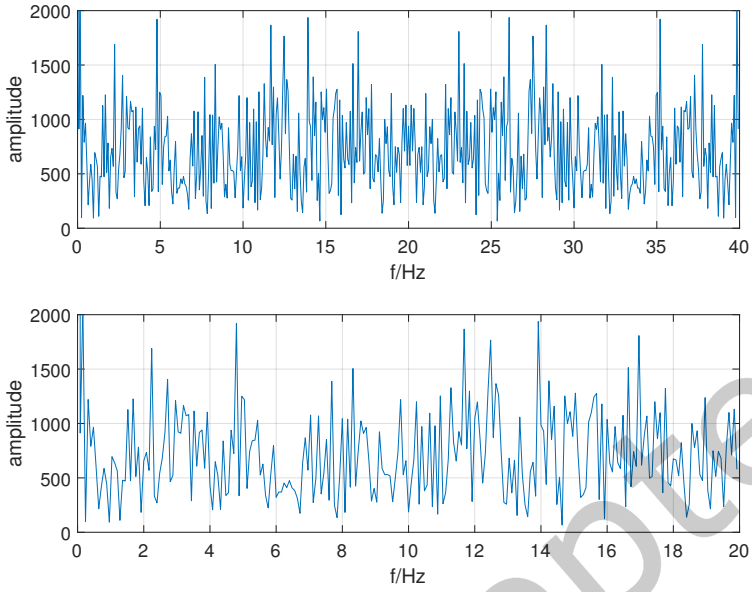


Fig. 5. Harmonic analysis Field

processing. It is the Fourier Series Decomposition of periodic signal to get the component whose frequency is an integral multiple of the fundamental frequency. Harmonics are unnecessary components in signal processing, and their existence will lead to signal distortion and loss. Therefore, in order to find harmonics in time, the system uses Fourier Series Decomposition to judge whether harmonics have been produced. The specific process is as follows:

Given a function of period T , it can be expressed as an infinite series :(j is an imaginary units)

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k \cdot e^{jk\left(\frac{2\pi}{T}\right)t} \tag{23}$$

Where, it can be calculated as follows:

$$a_k = \frac{1}{T}\int_T x(t) \cdot e^{-jk\left(\frac{2\pi}{T}\right)t} dt \tag{24}$$

Take note of:Is a function of period T, so periodic signals with different values have harmonic relations (that is, they all have a common period T). When k=0, the corresponding term in Eq.18 is called the DC component. When k=1, it has a fundamental frequency, which is called the first harmonic or fundamental wave. Similarly, there are: second harmonic, third harmonic and so on. As shown in Fig. 5, there are two periods of the second harmonic component in the picture above and one period of the first harmonic component in the figure below. The generation of this harmonic is due to the simultaneous operation of the two motors, and the vibration frequency reaches $40Hz$ during operation, resulting in resonance phenomenon at a certain time.

The generalized Fourier series is similar to the orthogonal decomposition of vectors in geometric space, and the Fourier series of periodic functions is the orthogonal decomposition of functions in inner product space. Its

orthogonal decomposition is extended from $\left\{e_k = e^{ikx}, k \in \phi\right\}$ basis to Legendre (1775-1837) polynomial and Haar (1885-1993) wavelet basis, etc.

For any system of orthogonal functions $\{\phi(x)\}$, if the function $f(x)$ defined on has only a finite number of discontinuity points of the first kind, then if $f(x)$ satisfies the closure equation:

$$\int_a^b f^2(x)\,dx = \sum_{k=1}^{\infty} c_k^2 \tag{25}$$

The series

$$\sum_{k=1}^{\infty} c_k \phi_k(x) \tag{26}$$

Must converge to $f(x)$, where:

$$c_n = \int_a^b f(x)\,\phi_n(x)\,dx \tag{27}$$

In fact, whether time converges or not, we always have:

$$\int_a^b f^2(x)\,dx \geq \sum_{k=1}^{\infty} c_k^2 \tag{28}$$

That's true. This is called Bessel's inequality. In addition, Eq.19 is easy to follow from orthogonality, since the projection on any unit orthogonal basis vector x is always $\langle x, e_i \rangle$.

Discover harmonics and deal with them in time to prevent the breakdown of the local power system due to excessive harmonics and prevent harmonics from causing parallel resonance or series resonance, amplifying harmonics, and causing capacitors to burn. To avoid harm caused by harmonics, an experimental model with multiple nonlinear components can be built by using a MATLAB environment in a multi-harmonic system. Ideally, his switch is a time pulse signal control. In the simulation experiment, it is concluded that the time when the harmonic source enters the circuit system affects the harmonics in the system, and the harmonics can achieve the effect of canceling each other.

Through Fourier Series Decomposition, this system can detect whether harmonics will be produced when multiple objects vibrate at the same time. The existence of harmonics affects the identity authentication of objects. It is essential to eliminate the influence of harmonics and accurately decompose the fingerprints of each object. The traditional signal analysis method is mainly the Fourier transform process. Fourier analysis is only suitable for stationary random signal analysis, and the entropy of the signal increases after study, and the correlation of the signal cannot be obtained. However, radio frequency signal belongs to non-stationary signal, and the traditional Fourier analysis method can not achieve good results. However, not only will not lose the Wigner-Ville transform the amplitude and phase information of the signal but also it has a clear concept of instantaneous frequency and group delay. This method has a high resolution and time-frequency aggregation, so it is suitable for analyzing non-stationary label signals. In this paper, the Wigner-Ville transform is used to process phase signals. Since it does not contain any window function, the contradiction between time resolution and frequency resolution cannot be considered in the linear time-frequency analysis method is avoided.

The non-stationary phase signal is processed by Wigner-Ville transform:

$$W_X(t,v) = \int_{-\infty}^{+\infty} X(v+\varepsilon/2)X^*(v-\varepsilon/2)e^{2\pi\varepsilon t}d\xi \tag{29}$$

The spectrum formula of the analytical signal of Wigner-Ville distribution is as follows:

$$W_x(t,v) = \int_{-\infty}^{+\infty} Z(v+\frac{f}{2}) Z*(v-\frac{f}{2}) e^{j2\pi\tau f} df \tag{30}$$

After that, the Wigner-Ville transform processing results are further converted, and the time-frequency domain analysis diagram is saved, the time-frequency diagram of mixed signals collected when multiple objects vibrate. At the same time, the phase data of a single thing during operation is also converted into a time-frequency graph, which is used as input data in the CNN model of convolutional neural network. Then, different types of phase signal data are eliminated one by one in the training process to reduce the complexity of subsequent model training until the fingerprint information of each object is matched one by one.

4.5 Separation of vibration signals of multiple objects

This topic intends to use the Markov Transition Field to optimize the complexity of the extracted one-dimensional temporal signals into two-dimensional image features to improve efficiency. For commercial reader receives, three signal indexes were analyzed, the RSSI values for position sensitivity and vulnerable to environmental impact, the Doppler value in the reader antenna when significant changes have taken place with the relative speed between tags will change, the phase values related to size and direction of the alternating current (ac), current frequency is higher, the greater the phase values. In the actual scenario, the relative moving speed of the reader antenna and the tag is small, so it is difficult for the commercial reader to capture such subtle changes. Therefore, in this topic, the phase information of tags is intensely mined and abstracted as a time series classification problem.

The tag phase information received by the reader antenna is a random variable, which is arranged forward and backward in time order. The distribution property at time t+1 is independent of the random variables before time T, and this property is consistent with the Markov property.Assuming time series $T = \{t_1, t_2, \cdots, t_N\}$, this topic divides the range into Q series, and each t_i in the time series information is mapped to the corresponding q_i . Through continuous mapping, a matrix W of $Q * Q$ can be obtained. $W_i j$ represents an overview of j a sequence of elements followed by i elements in the sequence, that is, $w_{ij} = P(at \in q_i | at-1 \in q_i)$ and satisfies $\sum Q_j = w_{ij} = 1$. So you get matrix $W = (w_{ij}) \leq i$ .Finally, the Markov Transition Field M is constructed. The Markov Transition Field M is the visualization matrix of the time series, namely the two-dimensional image feature. The matrix size is $[Q * Q]$, where the value of $M[i, j]$ is $M\left[X_i, X_j\right]$. The specific expression is as follows:

$$M = \begin{pmatrix} w_{\mathrm{i,j}}|x_1 \in X_i, x_1 \in X_j & w_{\mathrm{i,j}}|x_1 \in X_i, x_2 \in X_j & \cdots & w_{\mathrm{i,j}}|x_1 \in X_i, x_n \in X_j \\ w_{\mathrm{i,j}}|x_2 \in X_i, x_1 \in X_j & w_{\mathrm{i,j}}|x_2 \in X_i, x_2 \in X_j & \cdots & w_{\mathrm{i,j}}|x_2 \in X_i, x_n \in X_j \\ \vdots & \vdots & \ddots & \vdots \\ w_{\mathrm{i,j}}|x_n \in X_i, x_1 \in X_j & w_{\mathrm{i,j}}|x_n \in X_i, x_2 \in X_j & \cdots & w_{\mathrm{i,j}}|x_n \in X_i, x_n \in X_j \end{pmatrix} \tag{31}$$

The random field contains two essential characteristics: position and phase space. After assigning each position phase space in a random distribution, it only relates the value in any one part to its neighbors, not the others. Based on this feature, this paper converts the phase information of tags consistent with time series into Markov Random Field (Markov Random Field, MRF). It converts the coarse-grained phase features extracted by the reader into fine-grained image features with more exposed information and more specific elements. At the same time, the convolutional neural network is used to remove the phase features from the fine point of the Markov Field map. This method can not only recognize the fine features sensitively, but also it can overcome the defects of traditional signal processing. In the core of the convolutional neural network, features are extracted mainly through large convolutional nuclei. For example, VGG16, AlexNet, Resnet18, etc. The larger the size of the convolution kernel is, the stronger the ability to explore the information features will be, and more parameters will be introduced to improve the computational load and complexity. In this experiment, the Image Data Generator method is used to

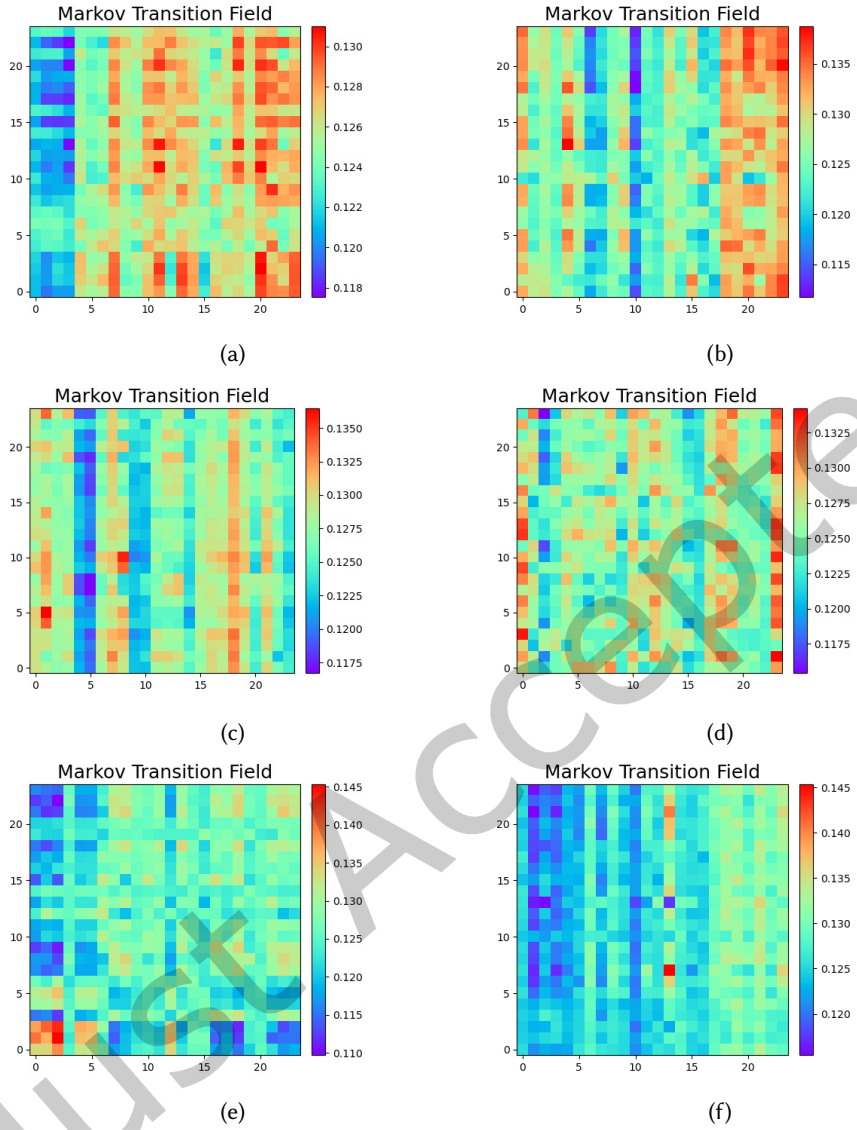


Fig. 6. Markov Transition Field

preprocess the image. In contrast, the idea is normalized, and the pixel value is reduced to between 0 and 1, which is conducive to the convergence of the model and to avoid the "death" of neurons. The phase MRF of regular rotation, reverse rotation, oil contamination penetration, and screw loosening are examples. These processed Markov diagrams constitute the data set of the subsequent fault classification model, which provides conditions for us to study signal characteristics and identify mechanical faults in the next step.

4.6 Fault Detection and Classification

In our method, a series of processing, such as feature amplification, harmonic detection, and cancellation, is carried out on the signals the commercial reader receives, making accurate matching of signal features and fault categories the key to the experiment. Therefore, we use neural networks to deeply learn the signal features to achieve precise fault category recognition.

In this system, the phase information of the tag is intensely mined and the fault classification and detection of the equipment is analyzed from the perspective of phase. The actual difference is π. Phase signal belongs to a one-dimensional time series signal. In this paper, the phase information of tags is abstracted as a time series classification problem. The tag phase information received by the reader antenna is a random variable arranged forward and backward in time order. The distribution property at time $t+1$ IS independent of the random variable before time, and this property is consistent with the Markov property.

From the above Fig. 6, it isn't easy to see the characteristic information of the specific phase signal in the picture. The data set composed of a Markov graph is put into the neural network to train, optimize the model and get the result. Deep neural networks can learn the input sample data layer by layer, especially by extracting the data features from the details of the picture information. Among them, Convolutional Neural Networks (CNN) are widely used in image recognition. CNN can overcome the defects of traditional signal processing and is particularly sensitive to recognizing small features. A convolutional neural network's core is extracting features by a large convolution kernel. For example, AlexNet, Resnet18, etc. The larger the size of the convolution kernel, the stronger the ability to explore information features. This way, more parameters will be introduced, improving the computation amount and complexity. VGG network uses 3×3 convolution kernels in each block to replace the large-size convolution kernels. Three 3×3 convolution kernels should be equivalent to the receptive fields of 7×7 convolution kernels. Still, the parameters are reduced by nearly half, which improves the model accuracy and reduces the computation simultaneously.

The original VGG model can have as many as 16 layers, including 13 convolutional layers and three fully connected layers, and can output 1000 classifications. The specific parameters are as follows: the convolution kernel size is 3×3, the step size is 1, and the convolution padding is 2. The size of the pooling layer is 2×2, and the step size is 2. There are two fully connected layers, and each layer contains 4096 neurons. The softmax layer is used as the output layer, and ReLU is used as the activation function. The VGG model is as follows Fig. 7.

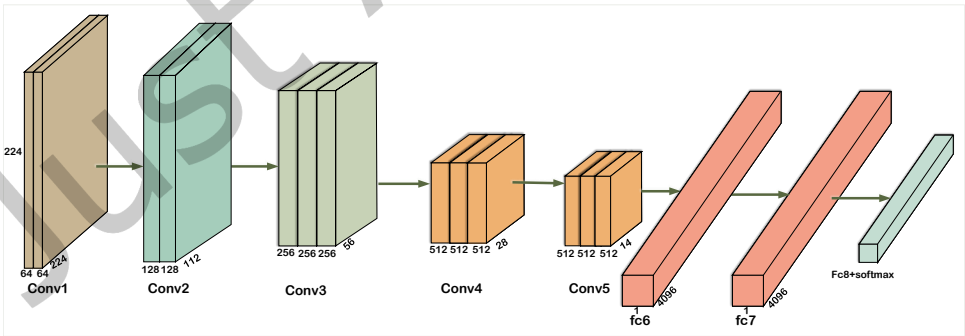


Fig. 7. Vgg16 network model

We use the Image Data Generator method to preprocess the image. The image in this paper is not cropped, rotated, or other operations to avoid errors due to the compression of image pixels. The original image is 270*270 pixels. After loading, the image is normalized, and the pixel value is reduced to between 0 and 1, which is

conducive to the convergence of the model, and the "death" of neurons is avoided. The images in this paper are composed of the vibration equipment's forward and reverse phase information, and most of the generated images are regular pixels. Moreover, the main objectives of this part belong to two categories, normal and abnormal. To improve the computational speed and reduce the model complexity, we simplify the VGG model step by step (as shown in Fig. 8).

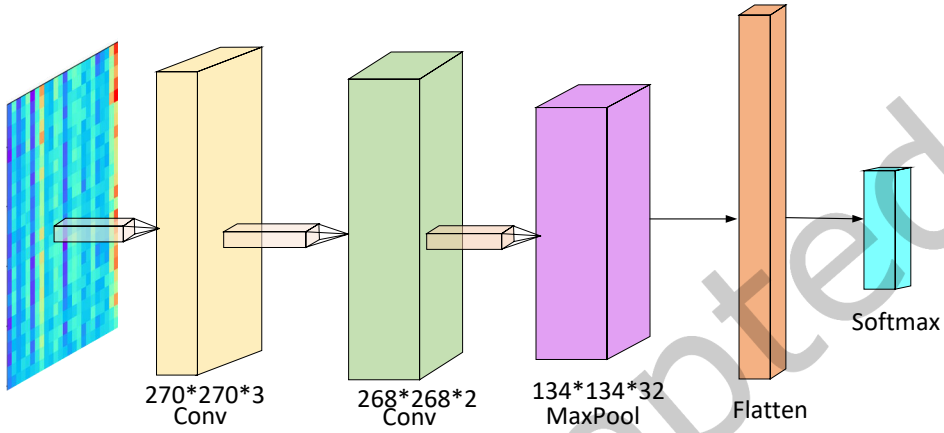


Fig. 8. Simplified VGG model

Convolutional layer: The function of the convolutional layer is to perceive some features of the picture locally. The convolution layer uses a rectangular convolution kernel to traverse each pixel on the image. We use a 3×3 convolution kernel for convolution operation, which can reduce the network parameters and ensure the network's sparsity. The image and convolution checkpoints and the bias parameter should be multiplied and summed. The value in the output feature map can be obtained. The feature map after convolutional calculation is expressed as follows:

$$x_{j+1} = W_j \otimes x_j + b_j \tag{32}$$

In the above equation is the current input feature map, which represents the convolution operation, means the weight of the convolution kernel, and represents bias.

Activation layer: The function of the activation layer is to realize the nonlinear transformation of the feature map after convolution operation through the activation function. The nonlinear modeling ability of the neural network is improved so that the neural network has the ability of hierarchical nonlinear mapping learning. Common activation functions include Sigmoid, Tanh and ReLU functions. For Eq.10, we can obtain after adding the activation function.

$$y_{j+1} = f(W_j \otimes x_j + b_j) \tag{33}$$

The convolution layer in this paper is convolved by a 3*3 convolution kernel and outputs 32 filters, which are finally activated by ReLU function. If x value is greater than 0, x is itself, otherwise x value is 0, which is used to solve the problem of nonlinear classification.

$$f(x) = (0, x) = \begin{cases} x, x > 0 \\ 0, x \leq 0 \end{cases} \tag{34}$$

Table 1. Comparison of five models

Model	Normal	Positive inversion	Greasy drit	Screw looses
SVM	71.85	84.56	80.19	71.92
Svgg	98.53	99.44	97.94	99.91
Resnet	82.66	93.82	85.04	80.52
Alexnet	83.06	94.38	85.93	83.92
Back propagation	86.03	71.53	74.27	76.82

Pooling layer: After the convolution operation, the model improves the ability to extract features, but the output dimension will also increase. The pooling layer can compress the sizes of each submatrix of the input tensor without losing the original main features. In this way, the calculation of parameters can be reduced, and the main representative features can be screened. The pooling mode selected in this model is maximum pooling, and the maximum value is taken for output to amplify the features further. The expression is as follows.

$$pooling(f\left[j-1\right], f\left[j\right], f\left[j+1\right]) = \max(f\left[j-1\right], f\left[j\right], f\left[j+1\right]) \tag{35}$$

Dropout layer: In order to prevent overfitting of model results, during the training process of the network model, we randomly discard neurons in the neural network according to a certain probability. In other words, these neurons are not involved in training.

Flatten layer: The multidimensional input data is one-dimensional, and we use it in the transition from the convolutional layer to the fully connected layer

Fully connected layer: The fully connected layer is the connection transition part between the convolutional network and the Softmax classifier. After the feature map passes through the Flatten layer, the convolutional network transforms the data into a one-dimensional array, and outputs the classification results through the fully connected layer.

Softmax classifier: Softmax is the commonly used last activation function. Its function is to convert the classification result into a probability value with a sum of 1. The function is as follows:

$$Soft\max(x)_i = \frac{\exp(x_i)}{\sum \exp(x_i)} \tag{36}$$

We summarize the proposed model. First, the input data is 270×270 training image, and after a 3×3 convolution kernel and 32 filters, the 268×268×32 dimensional matrix is output. Then, a 2×2 maximum pooling layer is carried out to output 134*134*32 dimensional data. After a series of feature extraction, the matrix is flattened by a Flatten layer and input to the fully connected layer with 32 neurons. Finally, the Softmax function is used to activate and output the two classifications, that is, the fault identification of vibration equipment is realized.

As shown in Tab. 1, five models are used to detect four types of faults, roughly divided into four categories: normal operation, motor reversal, screw loosening, and oil penetration. In this table, it can be seen that the accuracy of SVGG model is the highest among the four types of faults, and each accuracy is very average, indicating that the accuracy and stability of this model are high.

5 EXPERIMENT SETTINGS

5.1 System Deployment

In this paper, a non-contact vibration sensing and fault monitoring system based on a single tag are designed. The system consists of data acquisition, processing and analysis. The first part is to prepare for the experiment, which is mainly responsible for building a single-chip microcomputer to control the rotation speed of the motor

rotor. The second part is the feature extraction module, which is primarily responsible for selecting the best evaluation index and solving the problem that the vibration information is weak and difficult to handle. The third part is the core design, which mainly uses deep learning to realize motor frequency identification and fault classification detection. Next, we will detail the experimental deployment.

The system is used to detect faults in industrial systems, which we generally classify into four categories: regular operation, reversal, oil spills, and loose screws. A single tag and an antenna are deployed around the motor, and the reader reads the vibration information of the tag to realize the non-contact sensing system. In data preprocessing, we smoothed and de-noised the measured signal by a low-pass filter, extracted the key features from the ground signal, and converted the phase characteristic data with time series into a Markov Transition map using Markov Transform Field. Then, the deep learning model is used to train the Markov Transition graph and realize the fault classification and recognition.

5.2 Experimental Equipment

5.2.1 Hardware Devices. In the experiment, we firstly use ImpinJ Speedway R420 reader, which adopts EPC Gen2 standard and its working frequency band is between 920 and 925MHz. The reader mode is max throughput, so that the measured value can reach the maximum throughput. Secondly, in order to further maximize the number of observed samples, the tag of LLRP protocol is used to communicate with the reader. Thirdly, it is an RFID directional antenna. Finally, it combines the single chip microcomputer and speed meter, so that the speed and steering of the motor can be easily controlled and measured. During the experiment, the reader can reflect the RSSI, Phase, Doppler and other information of the corresponding tag, and synchronously transmit it to the laptop for processing and monitoring.

5.2.2 Software Facilities. In this experiment, we used the embedded Impinj LLRP toolkit to communicate with the reader. Impinj Reader modified this protocol to support phase read reporting. As for the client software,we used C to realize the network connection, MATALB to realize the signal processing and PYTHON to realize thefault classification. In the overall experiment, we are equipped with Intel i5-8265u CPU and 8GBRAM Lenovo PC, which makes the software implementation compatible with LLRP toolkit.It is more simple and direct to get various types of readings from the RFID reader.

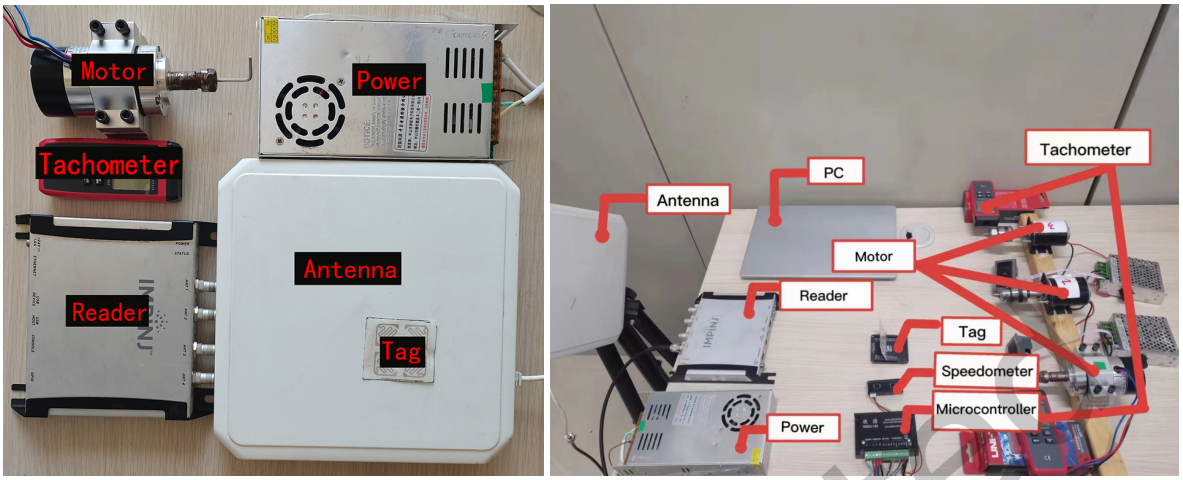
5.2.3 Experiment scene. These devices take the frequency data converted by manual motor speed adjustment as the basis for a series of explorations. The instrument deployment includes an R420 reader, directional antenna, tag, motor, microcontroller, tachometer, speedometer, Lenovo PC and other devices (as shown in Fig. 9) that enable real-time vibration sensing and identification. The deployment of the actual scene not only changes the conventional mode but also lays the foundation for the realization of efficient identification industrial equipment vibration sensing.

In this experiment, the maximum sampling rate of 140Hz was achieved. Still, the distance between tag and reader based on it needed to be bigger, which was contradictory to the actual situation. To be more consistent with reality, we selected the sampling rate 60Hz with the best practical effect for the experiment.

6 SYSTEM EVALUATION

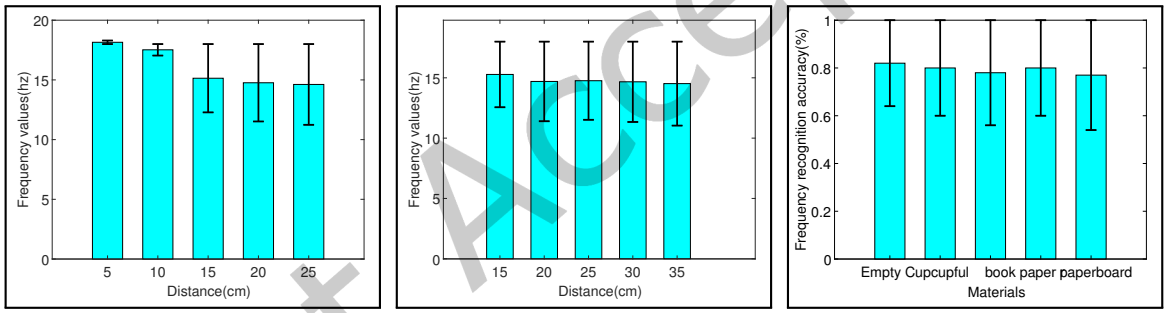
6.1 System analysis

In this paper, a single-tag multi-dimensional non-contact fault identification and detection system is designed, and the Markov Transition Field amplifies the signal features to make up for the single-tag design of the system; the non-contact sensing method further improves the accuracy problem and reduces the. A series of unnecessary costs due to contact and errors in experimental results caused by contact. This experiment achieved a maximum



(a) Diagram of experimental equipment (b) Experiment layout diagram

Fig. 9. System experiment equipment and deployment diagram.



(a) Distance between the device and the tag (b) Distance between the antenna and the tag (c) Impact of obstacles

Fig. 10. Impact of distacne and blocking items.

sampling of 140Hz, but the distance between the reader itself and the tag should not be too large, which contradicts the test in the natural environment.

In this test, we manually changed the speed of the equipment from 600 to 8400 rpm, that is, 10 to 140Hz, and the rotation interval was 60 rpm. At each speed, we performed 50 tests to recover the signal and match it to the speed of each device and then studied the matching accuracy of the system at each rate. Through experiments, the matching accuracy of different vibration stages is obtained. It is evident from this that the higher the rotational speed, the higher the recognition-matching accuracy. It can be seen that compared with high-speed devices, fine-grained low-speed devices are more challenging to match and identify. Still, overall, the recognition accuracy has reached a high level, which can be universally applied to various experimental scenarios.

6.2 Distance effect

6.2.1 Distance between motor and tag. In this experiment, we changed the distance between the tag and the reader from 5cm to 25cm, and the step size was 5cm. Since the accuracy of $25Hz$ in frequency identification is relatively high, we keep the vibration frequency at $25Hz$, and the distance between the tag and the antenna is unchanged. We have conducted several tests on each space and selected the best one. Fig. 10a shows the relevant errors, from which we can observe that when the distance between the tag and the motor gradually increases, the corresponding recognition frequency has an increasing error from the original $25Hz$ setting. It can be concluded that when the distance between the tag and the motor is too large, the reflected signal will be too weak, significantly reducing the frequency and forward and reverse identification accuracy. To further enhance the accuracy and sensing range, we can select more effective directional antennas to concentrate the energy compared with before, improving the experimental accuracy and sensing range.

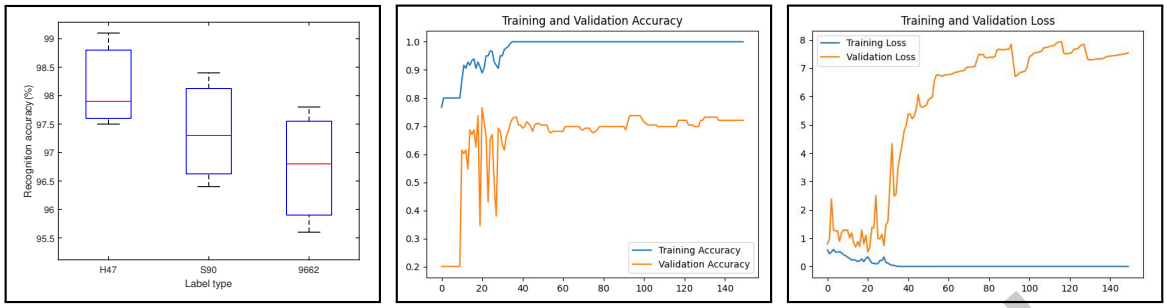
6.2.2 Distance between the antenna and the tag. In this experiment, we change the communication range of the COTS RFID hardware tag and antenna from 15cm to 35cm, with a step size of 5cm. The initial value of the selected range is 15cm to carry out the experiment in accordance with common sense and relatively accurate under the condition that the distance between the tag and the motor is unchanged. The same investigation was carried out with multiple measurements at each space, and it is evident from Fig. 10b that the farther the tag is from the antenna, the more the identified frequency magnitude differs from the actual $25Hz$ frequency. It can be concluded that the average error increases correspondingly when the distance increases. This is due to the more significant signal in the air propagating the space, indirectly weakening the recognition accuracy. To achieve the best recognition performance, we also suggest that the distance between the antenna and the tag should be kept in the middle of 15cm left and correct to effectively improve the accuracy of frequency detection under the condition of device placement.

6.3 Blocking effect

In this experiment, we also performed relevant operations in the challenging non-line-of-sight (NLOS), such as we use objects of different materials (paper, cardboard, carton, empty water cup, and total water cup) to block the signal transmission between the tag and the reader respectively, in which attention should be paid to keeping the distance between the tag, motor, and antenna constant during the experiment. Under this condition, it can be clearly seen from Fig. 10c that blocking has a high impact on the identification accuracy of frequency and the identification accuracy of forward and reverse, which significantly reduces the accuracy of frequency identification. However, even with the influence of blocking, the recognition accuracy is still around 0.8, and the error is within a specific range. This critical condition illustrates its universality and indirectly provides an excellent possibility for its deployment in the concrete operational environment.

6.4 Types of tags

In this experiment, we selected different types of tags for vibration monitoring and fault identification, including three commonly used tags: One is E51 with a size of 95*8mm, the other is AZ-9662 with a size of 70*17mm, the other E52 with a size of 68*14mm, and the last one is H47 with a size of 44*44mm. It can be seen from Fig. 11a that the identification accuracy of the H47 tag is higher. After consideration, the reason may be due to the smaller cross sections of other tags, which will weaken their coupling effect more than that of the H47 tag. However, it should be noted that the coupling effect will be affected by the change in distance of the tag reader. Therefore different types of tags can achieve more similar accuracy by significantly reducing the distance.



(a) Impact of tag type (b) Accuracy of iteration (c) Loss function iteration rate

Fig. 11. The impact of tag types and the number of model training iterations.

6.5 Number of iterations

Through the continuous adjustment and change of the training set and test set in the experiment, it can be intuitively seen from the Fig.11b and the Fig. 11c that the number of iterations has a greater impact on the accuracy and loss function, and the more iterations, the higher the fault identification accuracy. In the experiment, we conducted 140 rounds of iterative training, and at about 40 rounds, the accuracy and loss rate gradually tended to be stable with little fluctuation, and the accuracy rate was close to 1 and the loss rate was close to 0.7, which were generally stable.

6.6 Model accuracy

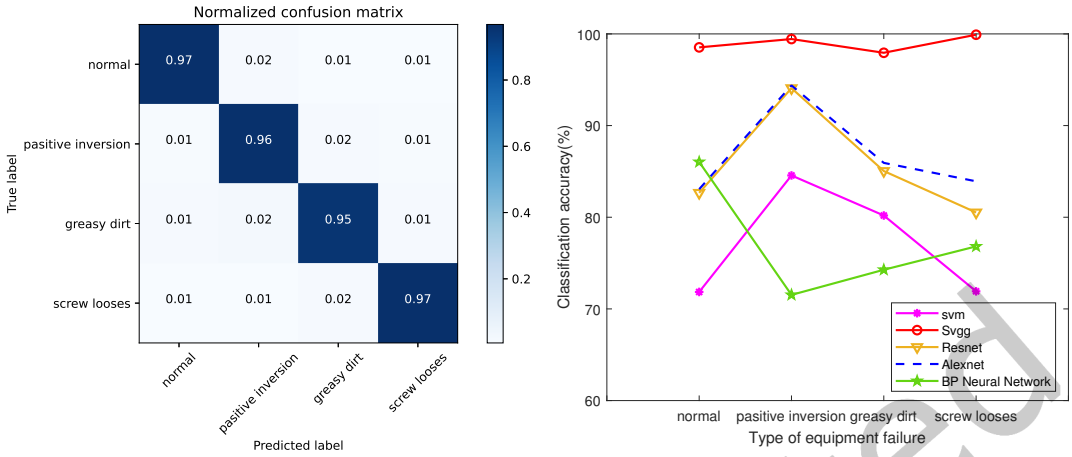
In the process of the experiment of fault identification and detection, the following two figures were used to analyze the accuracy and accuracy of the identification and detection. In the experiment, the fault data of the motor with $8Hz$ to $25Hz$ frequency is measured, and the Markov image formed by the phase Angle is classified and recognized by machine and even deep learning methods (SVM, Alexnet, Resnet, VGG16, BP neural network). Fig. 12 shows the types of classification detection models used and their detection accuracy. It is found that the improved SVGG model has better effect and accuracy compared with others.

7 CONCLUSION

This paper proposes a vibration sensing and fault identification detection system based on RFID single tag. The system realizes fine-grained non-contact vibration frequency sensing and real-time fault classification, detection, and recognition. The evaluation results show that the average accuracy of vibration frequency sensing and equipment fault detection and recognition can reach more than 95%, effectively breaking the traditional mode. Based on this, MFD can be applied to various devices in industrial systems.

ACKNOWLEDGMENTS

The research is supported by Intelligent Policing Key Laboratory of Sichuan Province (No. ZNJW2022KFZD004), Basic Research Plan of Shanxi Province (No. 202303021211339), Virtual Teaching and Research Office of Cyber Security (BJPC) of Ministry of Education (No. WAXVKF-2202), Anhui Natural Science Foundation (No. 2108085MF207), Shanxi Provincial Higher Education Teaching Reform and Innovation Project, Teaching Reform Project of Shanxi Police College.



(a) Confusion matrix of state identification

(b) Accuracy of different classification

Fig. 12. Accuracy of multiple state identification.

REFERENCES

[1] Jie Chen, Hengyu Guo, Guanlin Liu, Xue Wang, Yi Xi, Muhammad Sufyan Javed, and Chenguo Hu. 2017. A fully-packaged and robust hybridized generator for harvesting vertical rotation energy in broad frequency band and building up self-powered wireless systems. *Nano Energy* 33 (2017), 508–514.

[2] Anurag Choudhary, Deepam Goyal, S. L. Shimi, and Aparna Akula. 2018. Condition Monitoring and Fault Diagnosis of Induction Motors: A Review. *Archives of Computational Methods in Engineering* 26 (2018), 1221–1238.

[3] Lei Ding, Murtaza Ali, Sujeet Milind Patole, and Anand G. Dabak. 2016. Vibration parameter estimation using FMCW radar. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2016), 2224–2228.

[4] Debidatta Dwibedi, Yusuf Aytar, Jonathan Tompson, Pierre Sermanet, and Andrew Zisserman. 2020. Counting Out Time: Class Agnostic Video Repetition Counting in the Wild. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020), 10384–10393.

[5] Hyunjae Gil, Hyungki Son, Jin Ryong Kim, and Ian Oakley. 2018. Whiskers: Exploring the Use of Ultrasonic Haptic Cues on the Face. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018).

[6] Cihun-Siyong Alex Gong, Huang-Chang Lee, Yu-Chieh Chuang, Tien-Hua Li, Chih-Hui Simon Su, Lung-Hsien Huang, Chih-Wei Hsu, Yih-Shiou Hwang, Jiann-Der Lee, and Chih-Hsiung Chang. 2018. Design and Implementation of Acoustic Sensing System for Online Early Fault Detection in Industrial Fans. *J. Sensors* 2018 (2018), 4105208:1–4105208:15.

[7] Ayaz Kafeel, Sumair Aziz, Muhammad Awais, Muhammad Attique Khan, Kamran Afaq, Sahar Ahmed Idris, Hammam Alshazly, and Samih M Mostafa. 2021. An Expert System for Rotating Machine Fault Detection Using Vibration Signal Analysis. *Sensors* 21, 22 (2021), 7587.

[8] Harini Kolamunna, Thilini Dahanayaka, Junye Li, Suranga Seneviratne, Kanchana Thilakaratne, Albert Y. Zomaya, and Aruna Prasad Seneviratne. 2021. DronePrint: Acoustic Signatures for Open-set Drone Detection and Identification with Online Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5 (2021), 20:1–20:31.

[9] Rui Li, Nan-hua Yu, Bin-bin Zhang, Bing-kun Xu, Yi-dan Su, and Ming Gong. 2017. Research on algorithm of fault feature gene extraction based on time-frequency domain statistics. In *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. IEEE, 2390–2394.

[10] Yuxing Li, Bo Geng, and Shang bin Jiao. 2022. Dispersion entropy-based Lempel-Ziv complexity: A new metric for signal analysis. *Chaos, Solitons & Fractals* (2022).

[11] Anindya Maiti and Murtuza Jadliwala. 2018. Light Ears: Information Leakage via Smart Lights. *arXiv: Cryptography and Security* (2018).

[12] Julien NP Martel, Lorenz K Mueller, Stephen J Carey, Piotr Dudek, and Gordon Wetzstein. 2020. Neural sensors: Learning pixel exposures for HDR imaging and video compressive sensing with programmable sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42, 7 (2020), 1642–1653.

[13] Wang Mengjiao, Tang Zhenhao, Zhao Bo, and Hu Yunfeng. 2022. Wind Turbine Bearing Fault Diagnosis Method Based on Multi-domain Feature Extraction. In *2022 4th International Conference on Power and Energy Technology (ICPET)*. IEEE, 413–418.
[14] Ziqiang Qu, Huafeng Liu, Hao Ouyang, Chenyuan Hu, and Liang Cheng Tu. 2020. A High-sensitivity Optical MEMS Accelerometer based on SOI Double-side Micromachining. *2020 IEEE Sensors* (2020), 1–4.
[15] Xingrong Song, Di Zhang, Haiguo Tang, Jiran Zhu, Jian Sun, and Bin Zhang. 2022. Grounding Fault Line Selection Method of Distribution Network Based on Time-Frequency Domain Energy Matrix. In *2022 7th International Conference on Power and Renewable Energy (ICPRE)*. IEEE, 440–446.
[16] Feng Tian, Chunhui Zhao, Haidong Fan, Weijian Zheng, and Youxian Sun. 2019. Fault Diagnosis Based on EEMD and Key Feature Representation with Separation of Stationary and Nonstationary Signals. In *2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*. IEEE, 29–34.
[17] Xiaohua Tong, Kuifeng Luan, Uwe Stilla, Zhen Ye, Yusheng Xu, Sa Gao, Huan Xie, Qian Du, Shijie Liu, Xiong Xu, and Sicong Liu. 2019. Image Registration With Fourier-Based Image Correlation: A Comprehensive Review of Developments and Applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12 (2019), 4062–4081.
[18] Binbin Xie, Jie Xiong, Xiaojiang Chen, and Dingyi Fang. 2020. Exploring commodity RFID for contactless sub-millimeter vibration sensing. *Proceedings of the 18th Conference on Embedded Networked Sensor Systems* (2020).
[19] Chenhan Xu, Zhengxiong Li, Hanbin Zhang, Aditya Singh Rathore, Huining Li, Chen Song, Kun Wang, and Wenyao Xu. 2019. WaveEar: Exploring a mmWave-based Noise-resistant Speech Sensing for Voice-User Interface. *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services* (2019).
[20] Run Xu and Zhiqiang Chen. 2020. Technological Analysis on Motor Stall and Its Perspective. *Electrical Science & Engineering* 2, 1 (2020).
[21] Xiaolong Xu, Xincheng Tian, and Lelai Zhou. 2018. A Robust Incremental-Quaternion-Based Angle and Axis Estimation Algorithm of a Single-Axis Rotation Using MARG Sensors. *IEEE Access* 6 (2018), 42605–42615.
[22] Xiangyu Xu, Jiadi Yu, Yingying Chen, Qin Hua, Yanmin Zhu, Yi-Chao Chen, and Minglu Li. 2020. TouchPass: towards behavior-irrelevant on-touch user authentication on smartphones leveraging vibrations. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. 1–13.
[23] Song Xue and Ian Howard. 2018. Torsional vibration signal analysis as a diagnostic tool for planetary gear fault detection. *Mechanical Systems and Signal Processing* 100 (2018), 706–728.
[24] Kaoru Yamashita, Hideyuki Murakami, Latpasamixay Chansomphou, and Masanori Okuyama. 2003. Ultrasonic array sensor using piezoelectric film on silicon diaphragm and its resonant-frequency tuning. *TRANSDUCERS '03. 12th International Conference on Solid-State Sensors, Actuators and Microsystems. Digest of Technical Papers (Cat. No.03TH8664)* 1 (2003), 939–942 vol.1.
[25] Panlong Yang, Yuanhao Feng, Jie Xiong, Ziyang Chen, and Xiangyang Li. 2020. RF-Ear: Contactless Multi-device Vibration Sensing and Identification Using COTS RFID. *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications* (2020), 297–306.
[26] Weiqiang Yao, Xiaoqing Gao, Shanshan Wang, and Yang Liu. 2019. Distribution High Impedance Fault Detection Using the Fault Signal Reconstruction Method. In *2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*. IEEE, 2573–2577.
[27] Huaidong Zhang, Xuemiao Xu, Guoqiang Han, and Shengfeng He. 2020. Context-Aware and Scale-Insensitive Temporal Repetition Counting. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020), 667–675.
[28] Jingcheng Zhao, Xinru Fu, Zongkai Yang, and Fengtong Xu. 2019. UAV detection and identification in the Internet of Things. *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)* (2019), 1499–1503.
[29] Chao Zheng, Ke Zhu, Susana Cardoso de Freitas, Jen-Yuan (James) Chang, J. E. Davies, Pete Eames, Paulo P. Freitas, Olga Kazakova, Cheolgi Kim, Chi Wah Leung, Sy-Hwang Liou, Alexey V Ognev, S. N. Piramanayagam, Pavel Ripka, Alexander S. Samardak, Kwang Ho Shin, S. Y. Tong, M. J. Tung, Shan X. Wang, Songsheng Xue, Xiaolu Yin, and Philip W. T. Pong. 2019. Magnetoresistive Sensor Development Roadmap (Non-Recording Applications). *IEEE Transactions on Magnetics* 55 (2019), 1–30.