IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 14, 2020, accepted December 24, 2020, date of publication December 30, 2020, date of current version January 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3048000

A Novel Method for Predicting Fault Labels of Roller Bearing by Generalized Laplacian Matrix

JIAWEI GU[1], YANXUE WANG[1,2], (Member, IEEE), CHAOFAN HU[3], (Member, IEEE) AND ZEXI LUO[4]

[1]Beijing Key Laboratory of Performance Guarantee on Urban Rail Transit Vehicles, Beijing University of Civil Engineering and Architecture, Beijing 100044, China
[2]School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China
[3]School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin 541004, China
[4]Aviation Key Laboratory of Science and Technology on Fault Diagnosis and Health Management, AVIC Shanghai Aero Measurement Controlling Research Institute, Shanghai 200080, China

Corresponding author: Yanxue Wang (yan.xue.wang@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51875032, in part by the Fundamental Research Funds for Beijing University of Civil Engineering and Architecture under Grant X20159 and Grant X20061, in part by the Doctoral Research Foundation under Grant UF20027Y, in part by the Project of Director of Guangxi Manufacturing System and Advanced Manufacturing Technology Key Laboratory under Grant PF20105P, and in part by the Innovation Project of GUET Graduate Education under Grant 2020YCXS014.

ABSTRACT Because mechanical failures are accompanied by contingency and randomness, fault data is often difficult to obtain, and fault labels are also difficult to assign. The lack of data and fault labels have become important issues that restrict the development of fault diagnosis. The paper proposed a generalized Laplacian label prediction (GLLP) algorithm, which mainly uses the generalized Laplacian matrix and calculated a new locally smooth term. Therefore, data points with ambiguous and unclear labels will be assigned a small label value, while samples with more certain labels can get a more confident label value. The effectiveness of the method is verified on the public dataset and the real test rig dataset, and it is expected that this method can be extended to more complex mechanical system fault diagnosis.

INDEX TERMS Fault diagnosis, rotating machinery, laplacian matrix, label prediction.

I. INTRODUCTION

Mechanical systems are similar to medical systems. When faults occur in the system, it is vital to quickly and accurately find the points of faults and the causes of the faults. In actual mechanical systems, the collected fault data often lacks a large amount of label information, which is a fatal blow to most existing fault diagnosis methods based on supervised learning and deep learning [1], [2]. How to accurately classify the fault type using the only incomplete label at hand? This problem is a dilemma that current fault diagnosis research has to face. Based on this problem, a series of fault diagnosis methods based on few shot learning came into being [3], [4].

Fault diagnosis methods based on unsupervised learning are adopted in [5]–[8]. The labels of the fault samples are discarded directly, and the samples are clustered with the same category, so as to realize the fault diagnosis research in the unlabeled condition. This way does effectively solve the problem of insufficient sample labels, but the problem of the failure type cannot be determined. The intelligent fault diagnosis method based on unsupervised learning can only keep the distance of the samples of the same category closer, but cannot identify the fault category. In response to this shortcoming, [9]–[12] proposed a series of fault diagnosis methods based on transfer learning. These fault diagnosis methods can effectively solve the dilemma of the lack of data and labels by transferring the knowledge learned in other fields to the field of fault diagnosis, which is a research focus of intelligent fault diagnosis. However, how to judge whether the knowledge learned from other fields is relavant to the features of the existing data is a practical problem that transfer learning has to face. In order to solve this problem, the researchers returned their focus to the existing data and tried to find an effective solution from the data with missing labels, as shown in Fig.1. [13], [13]–[18] Proposed intelligent fault diagnosis methods based on semi-supervised learning, which assign the same labels with similar features by buiding similarity matrix. However, it is always very vague and ambiguous for the existing fault diagnosis methods based on semi-supervised learning to consider paired smooth items as the basis for constructing similarity matrices. The fault diagnosis method based on semi-supervised learning is always not

The associate editor coordinating the review of this manuscript and approving it for publication was Dazhong Ma.

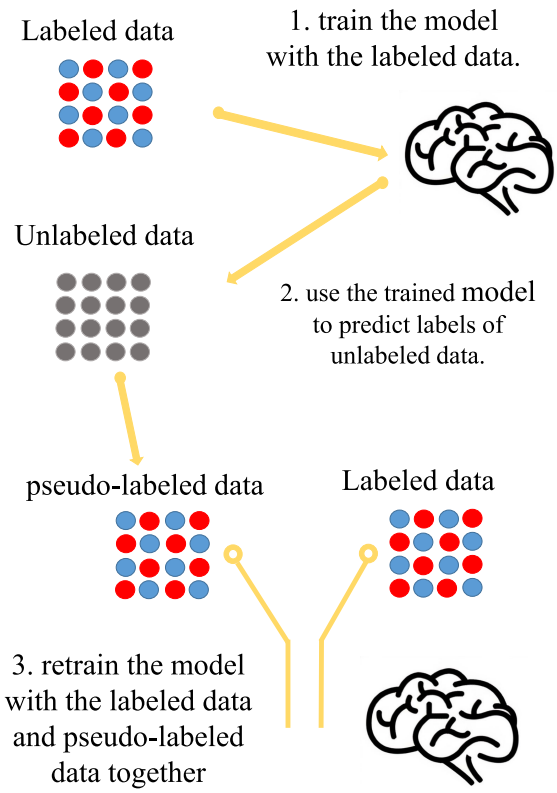


FIGURE 1. Principles of label propagation algorithm.

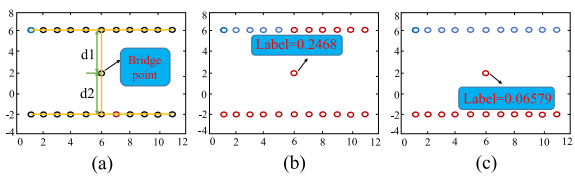


FIGURE 2. Display of smooth constraint items on artificial data sets.

satisfactory when propagating label information, as shown in Fig.2. In Fig.2(a), the blue, red, and black circles represent positive samples, negative samples and unlabeled samples, respectively. The circle with coordinates (6, 2) has the same distance to the positive and negative samples, so it can be divided into any category, this point is called the'ambiguity point'. Fig.2(b) describes that the ambiguity points are misclassified when only pairwise smooth terms are adopted. In response to this ambiguity, a generalized Laplacian matrix [19] is proposed to define a new smooth term. It can be seen from Fig.2(c) that the proposed new smooth term effectively prevents the label information from passing ambiguity points and achieves greater classification effect.

The main contributions of this research can be summarized as follows:

1) Solving the dilemma of insufficient data labels for fault diagnosis;
2) A new smooth term is constructed by adopting generalized Laplace matrix;
3) The proposed GLLP can be regarded as a unified framework for graph-based label propagation methods.

The rest of the paper is summarized as follows: The inference model and inductive model are detailed described in Section II. In section III, the qualities of label propagation and fault diagnosis will be verified on Double Moon dataset and MFS-MG test rig dataset. Some conclusions are drawn in Section IV.

II. THE PROPOSED METHOD

The proposed GLLP algorithm is divided into two parts: an inference model and an induction model. The inference model is drawn in Euclidean space, and the induction problem is described in Hilbert Regenerative Kernel space.

A. INFERENCE MODEL

The graph $\mathcal{W} = \langle \mathcal{Q}, \mathcal{R} \rangle$ is given. Existing graph inference algorithms [20], [21] usually use the original Laplacian matrix $\mathbf{H} = \mathbf{Z} - \mathbf{P}$ to describe the smoothness between labels. When the vector $\mathbf{u} = (u_1, u_2, \cdots, u_k)^\top$ is used to record the soft labels of all samples $\{c_a\}_{a=1}^k$ in ϕ, then the existing smooth terms can be described as:

$$\frac{1}{2}\sum_{a=1}^{k}\sum_{b=1}^{k}\theta_{ab}\,(u_a - u_b)^2 = \mathbf{u}^\top \mathbf{H}\mathbf{u}. \tag{1}$$

However, Fig.2 has shown that the paired smooth terms defined by Eq.1 cannot effectively handle ambiguous bridge points, so we redefine a new smooth term

$$\Gamma(\mathbf{u}) = \varpi \mathbf{u}^\top \mathbf{H}\mathbf{u} + \nu \mathbf{u}^\top (\mathbf{E} - \mathbf{Z}/r)\mathbf{u}, \tag{2}$$

among them, ϖ and ν are non-negative parameters used to impose different weights, $r = \sum_{a=1}^{k} l_{aa}$ represents the volume of graph $\mathcal{W}$. The second term on the right side of the formula is the locally smooth term, which can be further elaborated as

$$\mathbf{u}^\top (\mathbf{E} - \mathbf{Z}/r)\mathbf{u} = \sum_{a=1}^{k}(1 - l_{ab}/r)\,\mathbf{u}_a^2. \tag{3}$$

In the K-Nearest Neighbor graph $\mathcal{W}$, l_{aa} records the closeness of C_a and surrounding elements, so when Eq.3 is minimized, l_{aa} with a larger value can obtain a confidence label u_a, and a smaller sample of l_{aa} will obtain a relatively weak label.

The generalized Laplacian matrix expressed as $\widehat{\mathbf{H}} = \mathbf{E} - \beta\mathbf{P} - \beta^2(\mathbf{E} - \mathbf{Z})$ firstly appeared in [19], [22], where β is a variable parameter. When β is set to 1, the generalized Laplacian matrix $\widehat{\mathbf{H}}$ will be transformes into traditional Laplacian matrix, so the traditional Laplacian matrix can be regarded as a special case of our proposed method.

Let $\widehat{\mathbf{H}} = \varpi\mathbf{H} + \nu(\mathbf{E} - \mathbf{Z}/r)$, then Eq.2 can be expressed as $\Gamma(\mathbf{u}) = \mathbf{u}^\top \mathbf{H}\mathbf{u}$, where $\widehat{\mathbf{H}}$ and $\mathbf{H}$ have similar meanings in Eq.1. At the same time, further considering

$$\begin{aligned}\tilde{\mathbf{H}} &= \nu\mathbf{E} - \varpi\mathbf{P} + (\varpi - \nu/r)\mathbf{Z} \\ &= (\nu + \varpi - \nu/r)\left[\mathbf{E} - \frac{\varpi r}{\nu r + \varpi r - \nu}\mathbf{P}\right] \\ &\quad - (\nu + \varpi - \nu/r)\left[\frac{\varpi r - \nu}{\nu r + \varpi r - \nu}(\mathbf{E} - \mathbf{Z})\right],\end{aligned} \tag{4}$$

it can be found that when $\nu/\varpi = r(r-2)/r - 1$, and then divided by the coefficient $\nu + \varpi - \nu/r$, $\hat{\mathbf{H}}$ is the Eq.4.

Based on our proposed smoothing term, the inference model of GLLP in Euclidean space will be derived. Firstly, the initial state of the sample can be defined as $\mathrm{d} = (d_1, d_2, \cdots d_k)^\top$, when $d_a = 1$, c_a is a positive sample, when $d_a = 0$, c_a is an unlabeled sample. Then we extend the vector to the form of a matrix, and define a diagonal matrix $\mathrm{A}_{k\times k}$. if C_a is a labeled sample, the value of the i-th element on the diagonal element is 1, if it is an unlabeled sample, then set its value to 0. So the inference model of GLLP can be drawn as follows:

$$\min_{\mathbf{u}} G(\mathbf{u}) = \frac{1}{2}\left[\varpi \mathbf{u}^T \mathbf{H}\mathbf{u} + \nu \mathbf{u}^T(\mathbf{E} - \mathbf{Z}/r)\mathbf{u} + \|\mathbf{A}(\mathbf{u} - \mathbf{d})\|_2^2\right] \tag{5}$$

The first term in brackets in Eq.5 is the pairwise smooth term, which only considers the smoothness of the label between the two samples. Therefore, we implanted a local smoothing term, which can considers the sample and its nearby neighbor samples to achieve smoothness in the local area. In order to obtain the optimal solution of Eq.5, take the derivative of $G(\mathrm{u})$ to u and make the partial derivative equal to 0, then we can obtain

$$\varpi \mathbf{H}\mathbf{u} + \nu(\mathbf{E} - \mathbf{Z}/r)\mathbf{u} + \mathbf{A}\mathbf{u} - \mathbf{A}\mathbf{d} = \mathbf{0}. \tag{6}$$

Therefore, the form of the optimal solution of u can be described as:

$$\mathbf{u} = [\mathbf{A} + \varpi \mathbf{H} + \nu(\mathbf{E} - \mathbf{Z}/r)]^{-1}\mathbf{d}. \tag{7}$$

In Eq.7, if $u_a > 0$, then the label is defined as a positive label.

Theorem 1: Eq.5 is a convex optimization problem, and the optimal solution obtained is the global optimal solution.

Proof: The Hessian matrix of Eq.5 is drawn as follows:

$$\mathbf{O} = \mathbf{A} + \varpi \mathbf{H} + \nu(\mathbf{E} - \mathbf{Z}/r) \tag{8}$$

Through observation, it can be found that matrix $\mathbf{O}$ is diagonally dominant, so it is a positive definite matrix. In summary, theorem 1 is proved.

Next, two important parameters ϖ and ν of the proposed GLLP will be fully analyzed. The classification results of GLLP will prove to be insensitive to changes in parameters, and the result will be further proved in subsequent experiment. Based on Eq.8, the influence of ϖ and ν on U can be converted to study the stability of the solution of equation $\mathbf{O}\mathbf{u} = \mathbf{d}$. For this, a lemma is provided:

Lemma 2: A linear equation system $\mathbf{O}\mathbf{u} = \mathbf{d}$ is provided, assuming that $\mathbf{d}$ is certain, when the coefficient $\mathbf{O}$ has a slight disturbance $\iota\mathbf{O}$, then the relationship between the deviation of the result and the actual value $\mathbf{d}$ is as follows:

$$\frac{\|\iota\mathbf{u}\|}{\|\mathbf{u}\|} \le \frac{\mathrm{Cond}(\mathbf{O})(\|\iota\mathbf{O}\|/\|\mathbf{O}\|)}{1 - \mathrm{Cond}(\mathbf{O})(\|\iota\mathbf{O}\|/\|\mathbf{O}\|)}, \tag{9}$$

where $\mathrm{Cond}(\mathbf{O}) = \|\mathbf{O}\| \left\|\mathbf{O}^{-1}\right\|$ is the condition number.

1) STABILITY OF ν

When a slight disturbance appears on ν, $\iota\mathbf{O}$ in Eq.9 is $\iota\mathbf{O} = \iota\nu(\mathbf{E} - \mathbf{Z}/r)$, and the resulting error can be described as Eq.10, as shown at the bottom of the page, where $\zeta = 2\sum_{a=1}^{i}[\varpi l_{aa} + \nu(1 - l_{aa}/r)] + i > 0$. It can be found that in Eq.10, except for the different coefficients, the last term of the numerator and denominator is the same form, which shows that the disturbance has little effect on the result, so Eq.10 is infinitely close to 0.

2) STABILITY OF ϖ

When a slight disturbance appears on ϖ, $\iota\mathbf{O}$ in Eq.9 is $\iota\mathbf{O} = \iota\varpi\mathbf{H}$, and the resulting error can be described as Eq.11, as shown at the bottom of the page.

$$\begin{aligned}
\frac{\|\iota\mathbf{O}\|}{\|\mathbf{O}\|} &= \frac{\iota\nu\|\mathbf{E} - \mathbf{Z}/r\|}{\|\mathbf{A} + \varpi\mathbf{H} + \nu(\mathbf{E} - \mathbf{Z}/r)\|} \\
&= \frac{\iota\nu\sqrt{\sum_{a=1}^{k}(1 - l_{aa}/r)^2}}{\sqrt{\varpi^2\sum_{a=1}^{k}\sum_{b=1,b\neq a}^{k}\theta_{ab}^2 + \sum_{a=1}^{k}[\varpi l_{aa} + \nu(1 - l_{aa}/r)]^2 + \zeta}} \\
&= \frac{\iota\nu\sqrt{\sum_{a=1}^{k}(1 - l_{aa}/r)^2}}{\sqrt{\varpi^2\sum_a\sum_b\theta_{ab}^2 + \varpi\sum_a l_{aa}[(\varpi - 2\nu/r)l_{aa} + 2\nu] + \zeta + \nu^2\sum_a(1 - l_{aa}/r)^2}},
\end{aligned} \tag{10}$$

$$\begin{aligned}
\frac{\|\iota\mathbf{O}\|}{\|\mathbf{O}\|} &= \frac{\iota\varpi\|\mathbf{H}\|}{\|\mathbf{A} + \varpi\mathbf{H} + \nu(\mathbf{E} - \mathbf{Z}/r)\|} \\
&= \frac{\iota\varpi\sqrt{\sum_{a=1}^{k}l_{aa}^2 + \sum_{a=1}^{k}\sum_{b=1,j\neq a}^{k}\theta_{ab}^2}}{\sqrt{\varpi^2\sum_{a=1}^{k}\sum_{b=1,b\neq a}^{k}\theta_{ab}^2 + \sum_{a=1}^{k}[\varpi l_{aa} + \nu(1 - l_{aa}/r)]^2 + \zeta}} \\
&= \frac{\iota\varpi\sqrt{\sum_{a=1}^{k}l_{aa}^2 + \sum_{a=1}^{k}\sum_{b=1,j\neq a}^{k}\theta_{ab}^2}}{\sqrt{\nu^2\sum_a\left(1 - \frac{l_{aa}}{r}\right)^2 + \nu\sum_a\left\{\left(1 - \frac{l_{aa}}{r}\right)\left[\left(2\varpi - \frac{\nu}{r}\right)l_{aa} + \nu\right]\right\} + \zeta + \varpi^2\left(\sum_a l_{aa}^2 + \sum_a\sum_b\theta_{ab}^2\right)}},
\end{aligned} \tag{11}$$

Similar to the above proof, the final result of Eq.11 is very close to 0, Similar to the above proof, the final result of Eq.11 is very close to 0, which also shows that GLLP is not sensitive to changes in ϖ.

B. INDUCTIVE MODEL

The inductive model is established in the Hilbert space. Assuming that $F(\cdot,\cdot)$ is a Motzker kernel related to the Hilbert space and the corresponding norm is $\|\cdot\|_{\mathcal{H}}$, then the regularization expression of GLLP in the Hilbert space can be described as:

$$\min_{u\in\mathcal{H}_F} G(f) = \frac{1}{2}\left[\varsigma\|f\|_{\mathcal{H}}^2 + \varpi\mathbf{u}^\top\mathbf{H}\mathbf{u}\right] + \frac{1}{2}\left[\nu\mathbf{u}^\top(\mathbf{E}-\mathbf{Z}/r)\mathbf{u} + \sum_{a=1}^{k}(f(\mathrm{c}_a)-d_a)^2\right]. \tag{12}$$

Comparing Eq.12 with Eq.5, an inductive term $\|f\|_{\mathcal{H}}^2$ is added, which is used to control the complexity of the model and enhance the generalization ability of the model.

The literature [21] shows that the Eq.12 can be decomposed into a kernel function representing marked or unmarked:

$$u(\mathrm{c}) = \sum_{a=1}^{k} m_a F(\mathrm{c},\mathrm{c}_a). \tag{13}$$

When Eq.13 is substituted into Eq.12, a objective function can be described as:

$$\min_{\mathbf{M}\in\mathbb{R}^k} \widetilde{G}(\mathbf{M}) = \frac{1}{2}\left[\varsigma\mathbf{M}^\top\mathbf{F}\mathbf{M} + \varpi\mathbf{M}^\top\mathbf{F}\mathbf{H}\mathbf{F}\mathbf{M}\right] + \frac{1}{2}\left[\nu\mathbf{M}^\top\mathbf{F}(\mathbf{E}-\mathbf{Z}/r)\mathbf{F}\mathbf{M} + \|\mathrm{d}-\mathbf{A}\mathbf{F}\mathbf{M}\|^2\right], \tag{14}$$

where $\mathbf{F}$ is the Gram matrix defined on the training dataset, $\mathbf{F}_{ab} = F(\mathrm{c}_a,\mathrm{c}_b)$ $(1 \le a,b \le k)$. It is not difficult to find that Eq.14 is a convex function, so the model can obtain the global optimal solution, and the solution of Eq.14 can be obtained as

$$\mathbf{M} = [\varsigma\mathbf{E} + \varpi\mathbf{H}\mathbf{F} + \nu(\mathbf{E}-\mathbf{Z}/r)\mathbf{F} + \mathbf{A}\mathbf{F}]^{-1}\mathrm{d}. \tag{15}$$

1) ROBUSTNESS ANALYSIS

Definition 3: Suppose there is a metric space M_γ whose metric is γ, where W and $\hat{W} \subset M_\gamma$ are two sets in M_γ. If $\forall w \in W$, $\exists\hat{w} \in \hat{W}$, making $\gamma(w,\hat{w}) \le \upsilon$, then $\hat{W}$ is the υ -cover of W, and the υ of W is

$$N(\upsilon, W, \gamma) = \min\{|\hat{W}| : \hat{W} \text{ is an } \upsilon- \text{ cover of } W\}. \tag{16}$$

Definition 4 (Robustness): $\mathcal{G}$ represents the algorithm, ϕ and $H(\cdot)$ represent the training dataset and loss function of the algorithm respectively. If $\mathcal{Z}$ can be divided into η disjoint sets($\{N_a\}_{a=1}^{\eta}$), so that $\forall \mathrm{c}_1,\mathrm{c}_2 \in \phi$,

$$n_1, n_2 \in N_a \Rightarrow \left|H\left(\mathcal{G}_\phi, n_1\right) - H\left(\mathcal{G}_\phi, n_2\right)\right| \le \upsilon(\phi), \tag{17}$$

then $\mathcal{G}$ is called $(\eta, \upsilon(\phi))$ -robustness.

Based on Definition 3 and Definition 4, Definition 5 can be described as follows:

Definition 5: Input space is $\mathcal{C}$, and $\forall \mathrm{c}_a, \mathrm{c}_b \in \mathcal{C}$, $\|\mathrm{c}_a - \mathrm{c}_b\| \le \upsilon$. Based on the Gaussian kernel function, build a K-Nearest Neighbor graph $\theta_{ab} = \exp\left(-\|\mathrm{c}_a-\mathrm{c}_b\|^2/\left(2\omega^2\right)\right)$. When $N(\upsilon/2, \mathcal{C}, \|\cdot\|_2) < \infty$, the proposed GLLP is $\sqrt{\frac{8\,i}{\varsigma}}\left(1+\sqrt{\frac{i}{\varsigma}}\right)\sqrt{1-\exp\left(-\frac{\upsilon^2}{2\omega^2}\right)}$ -robustness.

Proof: when $\mathbf{M}$ in Eq.14 is set to $\mathbf{M}_0 = (0,\cdots,0)^\top$, then $\tilde{G}(\mathbf{M}_0) = \|\mathbf{d}\|^2/2 = i/2$ can be obtained. It can also be found that all the items in the brackets in Eq.14 are non-negative, $\frac{1}{2}\varsigma\mathbf{M}^\top\mathbf{F}\mathbf{M} \le G(\mathbf{M}) \le G(\mathbf{M}_0) = i/2$ can be further obtained, which means

$$\mathbf{M}^\top\mathbf{F}\mathbf{M} \le i/\varsigma \tag{18}$$

When faced with a binary classification problem, $\mathcal{Z}$ can be divided into $\eta = 2\,N(\upsilon/2, \mathcal{C}, \|\cdot\|_2)$ disjoint sets with an interval of υ [23]. According to Definition 3, if n_1 and n_2 are both a subset of set $N_a(1 \le a \le \eta)$, $\|\mathrm{c}_1 - \mathrm{c}_2\| \le \upsilon$ and $\|d_1 - d_2\| = 0$ can be obtained. The loss function of our proposed GLLP is

$$H(u) = (u(\mathrm{c}) - d)^2. \tag{19}$$

According to Definition 4, the difference of the loss function of the GLLP mapping function on sets n_1 and n_2 is

$$|H(u,n_1) - H(u,n_2)| = \left|(d_1 - u(\mathrm{c}_1))^2 - (d_2 - u(\mathrm{c}_2))^2\right|. \tag{20}$$

Taking Eq.13 into Eq.20, a specific loss function difference can be obtained as follows:

$$\begin{aligned}
&|H(u,n_1) - H(u,n_2)| \\
&= \left|\left[d_1 - \sum\nolimits_{a=1}^{k} m_a F(\mathrm{c}_1,\mathrm{c}_a)\right]^2 - \left[d_2 - \sum\nolimits_{a=1}^{k} m_a F(\mathrm{c}_2,\mathrm{c}_a)\right]^2\right| \\
&\le \left|d_1 + d_2 - \sum\nolimits_{a=1}^{k} m_a\left(F(\mathrm{c}_1,\mathrm{c}_a) + F(\mathrm{c}_2,\mathrm{c}_a)\right)\right| \\
&\quad d_1 - d_2 - \sum\nolimits_{a=1}^{k} m_a\left(F(\mathrm{c}_1,\mathrm{c}_a) - F(\mathrm{c}_2,\mathrm{c}_a)\right)| \\
&= |\mathbf{O}_1|\,|\mathbf{O}_2|,
\end{aligned} \tag{21}$$

where $\mathbf{O}_1 = d_1 + d_2 - \sum_{a=1}^{k} m_a\left(F(\mathrm{c}_1,\mathrm{c}_a) + F(\mathrm{c}_2,\mathrm{c}_a)\right)$, $\mathbf{O}_2 = d_1 - d_2 - \sum_{a=1}^{kk} m_a\left(F(\mathrm{c}_1,\mathrm{c}_a) - F(\mathrm{c}_2,\mathrm{c}_a)\right)$. The upper limit of $|\mathbf{O}_1|$ and $|\mathbf{O}_2|$ can be derived as follows:

$$\begin{aligned}
|\mathbf{O}_1| &\le |d_1| + |d_2| + |u(\mathrm{c}_1) + u(\mathrm{c}_2)| \\
&\le 2 + 2\max_{\mathrm{c}\in\{\mathrm{c}_1,\mathrm{c}_2\}} \langle u, F(\mathrm{c},\cdot)\rangle_{\mathcal{H}} \\
&\le 2 + 2\max_{\mathrm{c}\in\{\mathrm{c}_1,\mathrm{c}_2\}} \|u\|_{\mathcal{H}}\sqrt{F(\mathrm{c},\cdot)} \\
&\le 2 + 2\max_{\mathrm{c}\in\{\mathrm{c}_1,\mathrm{c}_2\}} \left\|\sum_{a=1}^{k} m_a F(\mathrm{c}_a,\cdot)\right\|_{\mathcal{H}} \sqrt{F(\mathrm{c},\cdot)} \\
&\le 2 + 2\sqrt{\left\langle \sum_{a=1}^{k} m_a F(\mathrm{c}_a,\cdot), \sum_{a=1}^{k} m_b F(\mathrm{c}_b,\cdot)\right\rangle_{\mathcal{H}}}
\end{aligned}$$

$$
\begin{aligned}
&= 2 + 2\sqrt{\sum_{a,b=1}^{k} m_a m_b F(\mathrm{c}_a, \cdot) F(\mathrm{c}_b, \cdot)} \\
&= 2 + 2\sqrt{\sum_{a,b=1}^{k} m_a F(\mathrm{c}_a, \mathrm{c}_b) m_b} \\
&= 2 + 2\sqrt{\mathrm{M}^\top \mathrm{FM}} \\
&\le 2 + 2\sqrt{\frac{i}{\varsigma}},
\end{aligned} \tag{22}
$$

where $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ represents the inner product in space $\mathcal{H}$.

In addition, since $\|u\|_{\mathcal{H}}^2 = \mathbf{M}^\top \mathbf{FM} \le \frac{i}{\varsigma}$, $\|u\|_{\mathcal{H}} \le \sqrt{\frac{i}{\varsigma}}$. At the same time, considering $\|d_1 - d_2\| = 0$, $\|\mathrm{c}_1 - \mathrm{c}_2\| \le \upsilon$, the upper limit of $\mathbf{O}_2$ can be obtained as follows

$$
\begin{aligned}
|\mathbf{O}_2| &= |u(\mathrm{c}_1) - u(\mathrm{c}_2)| \\
&= |\langle u, F(\mathrm{c}_1, \cdot) - F(\mathrm{c}_2, \cdot) \rangle_{\mathcal{H}}| \\
&\le \|u\|_{\mathcal{H}} \|F(\mathrm{c}_1, \cdot) - F(\mathrm{c}_2, \cdot)\|_{\mathcal{H}} \\
&= \|u\|_{\mathcal{H}} \sqrt{F(\mathrm{c}_1, \mathrm{c}_1) + F(\mathrm{c}_2, \mathrm{c}_2) - 2F(\mathrm{c}_1, \mathrm{c}_2)} \\
&\le \sqrt{i/\varsigma} \sqrt{F(\mathrm{c}_1, \mathrm{c}_1) + F(\mathrm{c}_2, \mathrm{c}_2) - 2F(\mathrm{c}_1, \mathrm{c}_2)} \\
&\le \sqrt{i/\varsigma} \sqrt{2 - 2\exp\left[-\|\mathrm{c}_1 - \mathrm{c}_2\|^2 / (2\omega^2)\right]} \\
&= \sqrt{i/\varsigma} \sqrt{2 - 2\exp\left[-\upsilon^2 / (2\omega^2)\right]}.
\end{aligned} \tag{23}
$$

Finally, taking Eq.22 and Eq.23 into Eq.21, the specific loss function difference can be obtained as follows:

$$
|H(u, n_1) - H(u, n_2)| \le \sqrt{\frac{8i}{\varsigma}} \left(1 + \sqrt{\frac{i}{\varsigma}}\right) \sqrt{1 - \exp\left(-\frac{\upsilon^2}{2\omega^2}\right)}, \tag{24}
$$

the proposed GLLP is $\sqrt{\frac{8i}{\varsigma}} \left(1 + \sqrt{\frac{i}{\varsigma}}\right) \sqrt{1 - \exp\left(-\frac{\upsilon^2}{2\omega^2}\right)}$ -robustness has been proved.

2) GENERALIZATION ABILITY ANALYSIS

Assuming that all samples are independent and identically distributed, then the empirical error and generalization error can be defined as $\tilde{H}(\mathcal{G}_\phi) = E_{\mathrm{c} \sim P}\left[H(\mathcal{G}_\phi, \mathrm{c})\right]$, $H_{emp}(\mathcal{G}_\phi) = \frac{1}{k} \sum_{\mathrm{c}_a \in \phi} H(\mathcal{G}_\phi, \mathrm{c}_a)$.

Definition 6: Assuming that there are n independent and identically distributed samples in the training dataset ϕ, and the algorithm $\mathcal{G}$ is robust, then for any $\iota > 0$ [24], there will be the following constraints when the probability is greater than or equal to $1 - \iota$:

$$
\left|\tilde{H}(\mathcal{G}_\phi) - H_{emp}(\mathcal{G}_\phi)\right| \le \upsilon(\phi) + U\sqrt{\frac{2\eta \ln 2 + 2\ln(1/\iota)}{k}}, \tag{25}
$$

where the upper bound of the loss function $H(\cdot, \cdot)$ is denoted as U.

Definition 7: Assuming that the loss function of the GLLP algorithm is $H(u, \phi) = (u(\mathrm{c}) - d)^2$, then for any $\iota > 0$, when the probability is greater than or equal to $1 - \iota$, the generalization error limit of GLLP is:

$$
\begin{aligned}
\left|\tilde{H}(G_\xi) - H_{emp}(G_\xi)\right| \le &\sqrt{\frac{8i}{\varsigma}} \left(1 + \sqrt{\frac{i}{\varsigma}}\right) \sqrt{1 - \exp\left(\frac{-\upsilon^2}{2\omega^2}\right)} \\
&+ 2\left(1 + \frac{i}{\varsigma}\right) \sqrt{\frac{2\eta \ln 2 + 2\ln(1/\iota)}{k}}
\end{aligned} \tag{26}
$$

Proof: The generalization error limit of GLLP is related to $\upsilon(\phi)$, η, and U. $\upsilon(\phi)$ and η have been obtained in the above part, and then only the upper bound U is required, so

$$
\begin{aligned}
H(u, \phi) &= (d - u(\mathrm{c}))^2 = d^2 - 2\,du(\mathrm{c}) + u^2(\mathrm{c}) \\
&\le 2\,d^2 + 2\,u^2(\mathrm{c}) \le 2 + 2i/\varsigma.
\end{aligned} \tag{27}
$$

so the upper bound of the loss function is

$$
U = 2 + 2i/\varsigma. \tag{28}
$$

Definition 7 can be proved by substituting Eq.24 and Eq.28 into Eq.25.

3) LINEARIZATION OF GLLP

The above inductive GLLP is nonlinear, but its corresponding linear model can also be sorted out.

According to Eq.13, the label of test sample c_0 is

$$
u_1(\mathrm{c}_0) = \mathrm{k}\mathbf{M} = \mathrm{k}\left[\varsigma\mathbf{E} + \varpi\mathbf{HF} + \nu\left(\mathbf{E} - \frac{1}{r}\mathbf{Z}\right)\mathbf{F} + \mathbf{AF}\right]^{-1} \mathrm{d}, \tag{29}
$$

where F is a kernel matrix.

Linear decision function $u(\mathrm{c}_0) = \theta^\top \mathrm{c}_0$ and data matrix $\mathbf{C} = (\mathrm{c}_1, \mathrm{c}_2, \cdots \mathrm{c}_k)$ are used to build the GLLP model:

$$
\begin{aligned}
&\min_\theta G(\theta) \\
&= \frac{1}{2}\left[\varsigma\|\theta\|^2 + \varpi\theta^\top \mathbf{CHC}^\top \theta\right] \\
&\quad + \frac{1}{2}\left[\nu\theta^\top \mathbf{C}\left(\mathbf{E} - \frac{1}{r}\mathbf{Z}\right)\mathbf{C}^\top \theta + \left\|\mathrm{d} - \mathbf{AC}^\top \theta\right\|_2^2\right],
\end{aligned} \tag{30}
$$

and its optimal solution can be obtained as:

$$
\theta = \left[\varsigma\mathbf{E} + \varpi\mathbf{CHC}^\top + \nu\mathbf{C}\left(\mathbf{E} - \frac{1}{r}\mathbf{Z}\right)\mathbf{C}^\top + \mathbf{CAC}^\top\right]^{-1} \mathbf{C}\mathrm{d}. \tag{31}
$$

The label of test sample c_0 is

$$
\begin{aligned}
&u_2(\mathrm{c}_0) \\
&= \mathrm{c}_0^\top \theta \\
&= \mathrm{c}_0^\top \left[\varsigma\mathbf{E} + \varpi\mathbf{CHC}^\top + \nu\mathbf{C}\left(\mathbf{E} - \frac{1}{r}\mathbf{Z}\right)\mathbf{C}^\top + \mathbf{CAC}^\top\right]^{-1} \mathbf{C}\mathrm{d}.
\end{aligned} \tag{32}
$$

Substituting $\mathbf{F} = \mathbf{C}^\top \mathbf{C}$ into Eq.29 can get

$$
u_1(\mathrm{c}_0) = \mathrm{c}_0^\top \mathbf{C}\left(\varsigma\mathbf{E} + \mathbf{UC}^\top \mathbf{C}\right)^{-1} \mathrm{d}, \tag{33}
$$

where $\mathbf{U} = \varpi\mathbf{H} + \nu\left(\mathbf{E} - \frac{1}{r}\mathbf{Z}\right) + \mathbf{A}$.

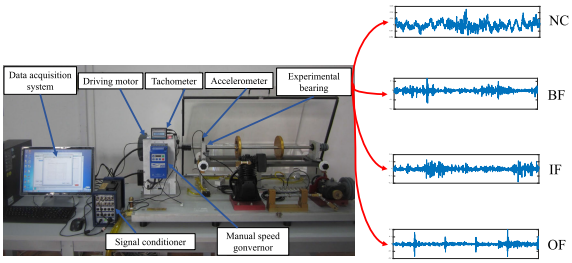


FIGURE 3. MFS-MG test rig.

According to the matrix inverse lemma, $u_1(c_0)$ and $u_2(c_0)$ can be derived as

$$\begin{aligned} u_1(c_0) &= c_0^\top \mathbf{C}\left(\varsigma \mathrm{E} + \mathrm{U}\mathrm{C}^\top \mathrm{C}\right)^{-1} \mathrm{d} \\ &= c_0^\top \mathbf{C}\left[\frac{1}{\varsigma}\mathbf{E} - \frac{1}{\varsigma^2}\mathbf{U}\left(\mathbf{E} + \frac{1}{\varsigma}\mathbf{C}^\top\mathbf{C}\mathbf{U}\right)^{-1}\mathbf{C}^\top\mathbf{C}\right]\mathrm{d} \\ &= c_0^\top \mathbf{C}\left[\frac{1}{\varsigma}\mathbf{E} - \frac{1}{\varsigma^2}\left(\mathbf{U}^{-1} + \frac{1}{\varsigma}\mathbf{C}^\top\mathbf{C}\right)^{-1}\mathbf{C}^\top\mathbf{C}\right]\mathrm{d}, \end{aligned} \tag{34}$$

$$\begin{aligned} u_2(c_0) &= c_0^\top \left(\varsigma \mathbf{E} + \mathbf{C}\mathbf{U}\mathbf{C}^\top\right)^{-1} \mathbf{C}\mathrm{d} \\ &= c_0^\top \left[\frac{1}{\varsigma}\mathbf{E} - \frac{1}{\varsigma^2}\mathbf{C}\left(\mathbf{M}^{-1} + \frac{1}{\varsigma}\mathbf{C}^\top\mathbf{C}\right)^{-1}\mathbf{C}^\top\right]\mathbf{C}\mathrm{d} \\ &= c_0^\top \mathbf{C}\left[\frac{1}{\varsigma}\mathbf{E} - \frac{1}{\varsigma^2}\left(\mathbf{U}^{-1} + \frac{1}{\varsigma}\mathbf{C}^\top\mathbf{C}\right)^{-1}\mathbf{C}^\top\mathbf{C}\right]\mathrm{d}, \end{aligned} \tag{35}$$

Eq.34 and Eq.35 are obviously equivalent.

III. EXPERIMENTAL ANALYSIS

The experimental analysis is divided into two parts: 1) Verify the label propagation ability of the GLLP algorithm on the artificial dataset; 2) Verify the fault classification ability of the proposed GLLP method on the dataset of MFS-MG test rig.

A. DESCRIPTION OF DATASETS

In this section, two types of datasets are adopted: Double Moon dataset and dataset on MFS-MG test rig. These two datasets will be introduced in detail as follows.

1) Double moon dataset: The dataset first appeared in [25]. 1000 samples are divided into two moon shapes in the section, the center coordinates of the two moons are (0,0), (10,0) respectively.
2) MFS-MG dataset: MFS-MG test rig is powered by a 1HP motor. The specific experimental equipment related parameters are presented in Table 1. The faults collected by the test rig include: inner fault (IF), outer fault (OF), roller fault (RF) and normal condition (NC). The specific structure of the MFS-MG test bench is shown in Fig.3.

B. VERIFICATION OF LABEL PROPAGATION PERFORMANCE

In the theoretical analysis part, the various performances of GLLP have been fully proved. Next, we will verify the label propagation ability of the GLLP method through experiments, and make the results more intuitively observed by everyone through the visualization method. In order to fully prove the superior performance of the proposed GLLP method, five latest graph-based label propagation

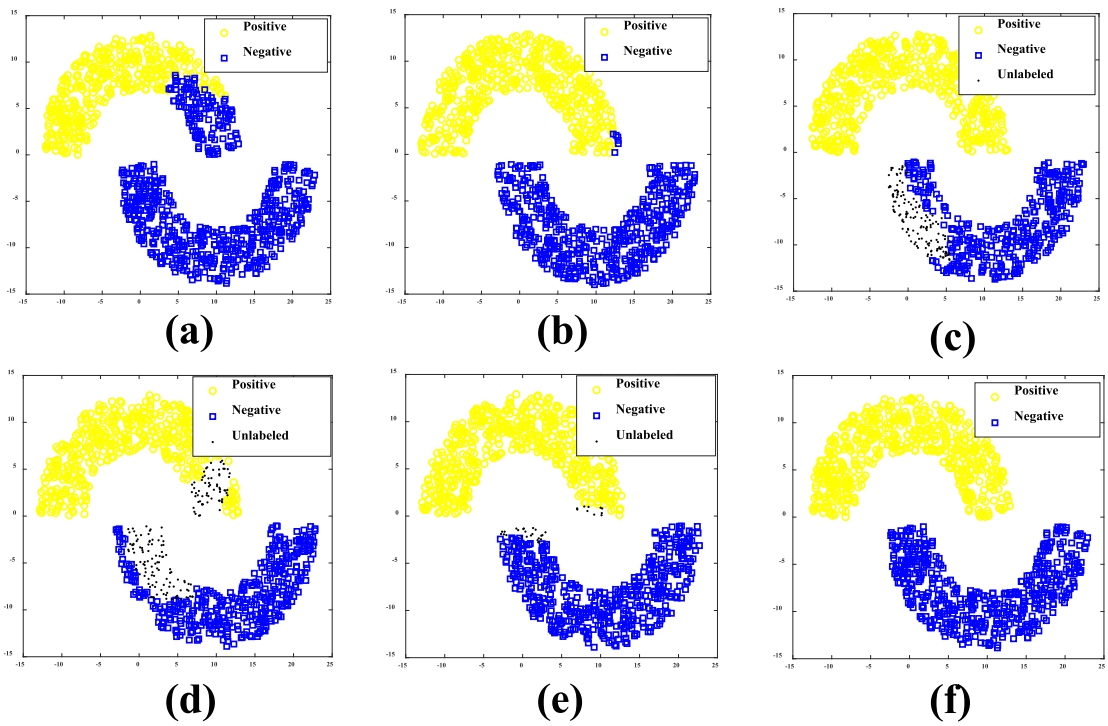


FIGURE 4. Comparison of label propagation on Doube-Moon Dataset.

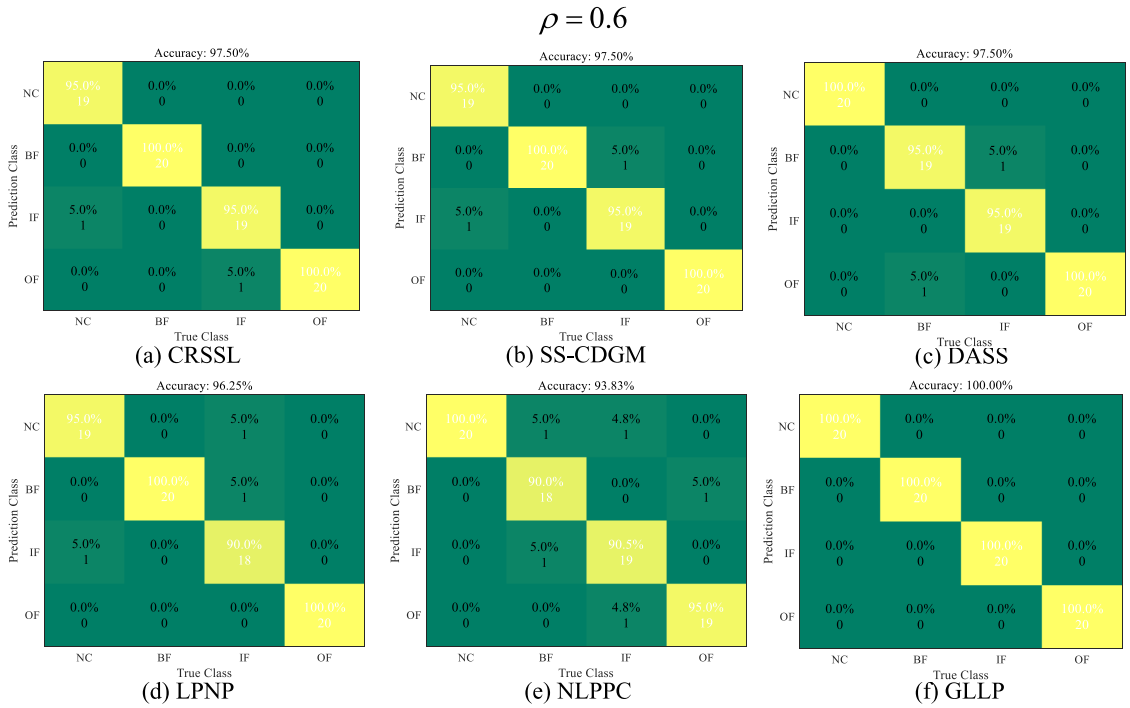


FIGURE 5. Comparison of the accuracy of label propagation when the sampling rate $\rho = 0.6$.

algorithms are selected as baselines for comparison: ALP [26] (auto-weighted label propagation):, BPFLP [27](belief-peaks clustering based on fuzzy label propagation), GLP [28](graph layout based label propagation), LNLP [29](linear neighborhood label propagation) and NLPPC [30](new label propagation algorithm with pairwise constraints). The parameters set by the six methods on Double Moon dataset are shown as:

1) ALP: $\alpha = 10^{-6}$, $\beta = 10^{-6}$, and $\gamma = 10^{-3}$;
2) BPFLP: $K = 2$ and $q = 0.7$;
3) GLP: $\alpha = 0.4$ and $\beta = 2$;
4) LNLP: $\alpha = 0.3$, $\beta = 0.7$, and $\rho = 1.0$;
5) NLPPC: $\mu = 0.6$ and $k = 2$;
6) GLLP: $\omega = 2$, $k = 5$, $\varpi = 1$ and $\nu = 10^{-3}$.

The results of the comparative experiment are presented in Fig.4. Since the GLLP algorithm we proposed is a non-iterative algorithm, we only give the final label propagation visualization results of all comparison methods. It can be easily seen from the results that the label propagation ability of our proposed GLLP method is the highest. In the process of label propagation, ALP and BPFLP propagates negative labels to positive labels, which is over propagation. After the propagation of LNLP, NLPPC and GLP, they still failed to completely spread all the labels, which is under propagation. The proposed GLLP method successfully propagated all kinds of labels to the correct position. The proposed GLLP can be considered as a state of the art label propagation method.

C. FAULT CLASSIFICATION OF MFS-MG TEST RIG

After the theoretical and experimental verification, the performance of the proposed GLLP method has been fully demonstrated. The GLLP method will be adopted for predicting the fault labels of rolling bearing and exploring the fault diagnosis performance.

In this work, 100 samples are selected to extract time-domain features, and each sample is a vibration sequence with a length of 1024.

In the experimental verification of the above subsection, five advanced algorithms are adopted for comparison. In this subsection, we adopt three state-of-the-art fault diagnosis methods based on semi-supervised learning for comparison. At the same time, to ensure the diversity of comparisons, two above latest label propagation algorithms are also adopted to compared with GLLP method. The parameters of these methods are set as follows: varsigma

1) CRSSL: Following the setting in [16];
2) SS-CDGM: Following the setting in [14];
3) DASS: Following the setting in [17];
4) LPNP: $\alpha = 0.2$, $\beta = 0.7$, and $\rho = 1.0$ [29];
5) NLPPC: $\mu = 0.5$ and $k = 4$ [30];
6) GLLP: $\omega = 2.5$, $k = 4$, $\varpi = 1$ and $\nu = 10^{-3}$.

Since the first five methods are all deep learning methods, the network level is complex, we only provide corresponding references, and the specific parameter settings will not be repeated. It is worth noting that in order to verify the quality of GLLP in the case of missing labels, the random sampling rate is set to $\rho = 0.2, 0.4, 0.6$. Under different random sampling rates, the prediction accuracy of the label propagation method is shown in Fig.5, Fig.6 and Fig.7. The ability of predicting labels are measured by confusion matrixs [31], [32]. In Fig.5, since the sampling rate adopted is 0.6, that is, the number of labeled samples is 60, all methods can learn enough knowledge and the accuracy rate is also the highest. In Fig.6,

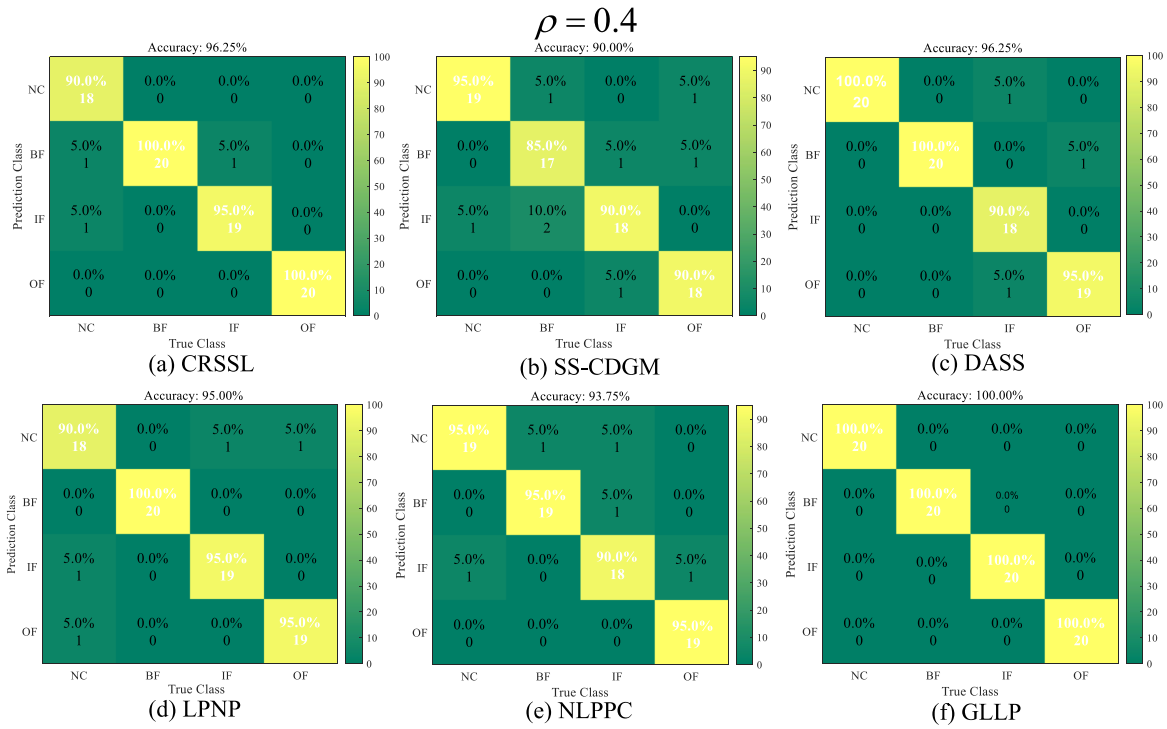


FIGURE 6. Comparison of the accuracy of label propagation when the sampling rate $\rho = 0.4$.

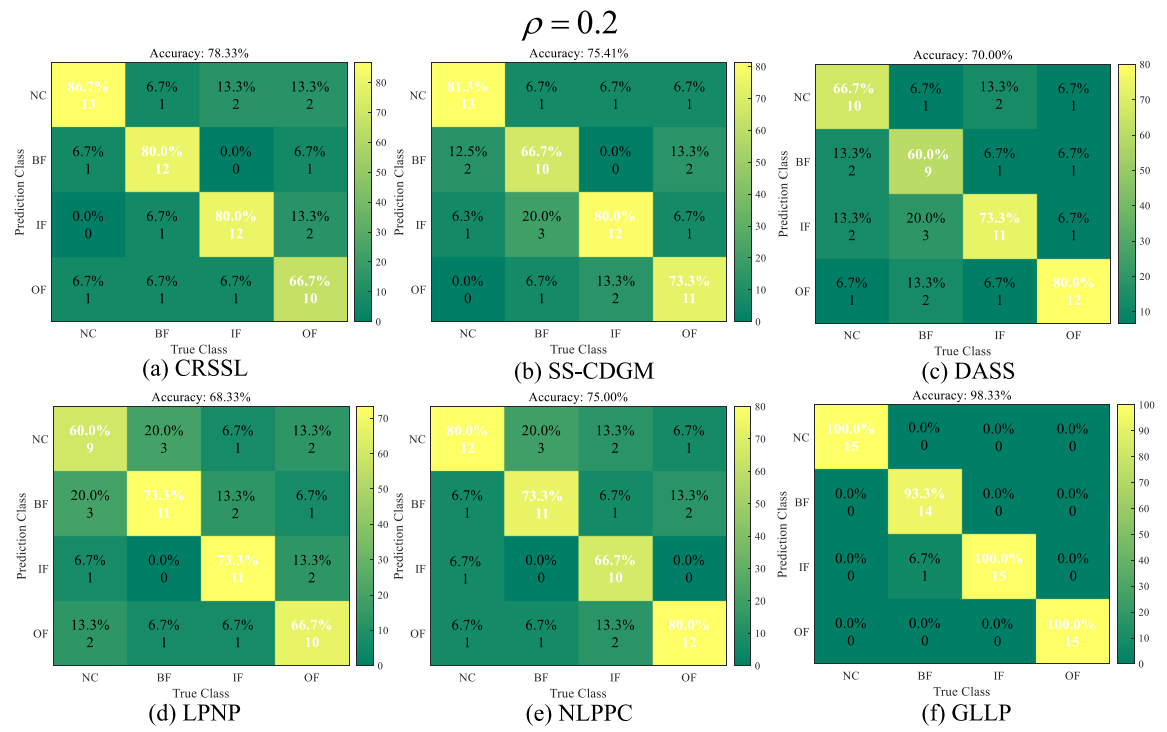


FIGURE 7. Comparison of the accuracy of label propagation when the sampling rate $\rho = 0.2$.

the accuracy of various methods for predicting labels is shown when the sampling rate is 0.4. It is shown in Fig.6 that when the sampling rate is 0.4, since the characteristics learned by various learning methods are sufficient, the accuracy of the predicted labels is generally high. and the GLLP method is the only method that completely predicts correctly. In Fig.7, the accuracy of various methods for predicting labels is shown when the sampling rate is 0.2. As shown in Fig.7, when the

sampling rate is 0.2, due to the lack of training samples, the accuracy of the comparison method predicting labels is all lower than 80%, but the accuracy of our proposed GLLP can still reach 98.33%, with only one label being predicted wrong.

IV. CONCLUSION

In this study, a novel label prediction method based on the generalized Laplacian matrix has been proposed. A new locally smooth term is constructed through the generalized Laplacian matrix to accurately determine the category of ambiguity points. The GLLP is a new attempt of the fault label prediction algorithm. The performance of the proposed GLLP method has been verified on theory, manifold artificial dataset and actual test rig dataset. In the process of label propagation ability verification, GLLP is the only method where label propagation is completely correct. In the verification process on the MFS-MG dataset, even in the extreme case where the sampling rate is only 0.4, the proposed GLLP method still has a correct rate of 98.33%. Therefore, it can be seen that the GLLP method has broad development prospect in actual industrial fault diagnosis.

In the future work, we will further study the application of the GLLP method in the problem of unbalanced data distribution, and will also explore the potential of GLLP in fault diagnosis of gears and non-rotating machinery.

ACKNOWLEDGMENT

The authors would like to thank the Editor, an Associate Editor and all reviewers who provided valuable comments on this paper. At the same time, the first author would like to thank his wife XiangXiang Yuan, in particular. During the study of the first author, Yuan took on all the work in the family and provided backing for the first author. Gu dedicates this work to Yuan and hopes that each Researcher who reads this paper in the future will realize that the completion of every work is inseparable from the people or things that support him silently. This work will be displayed as a poster at the wedding of Gu and Yuan.

REFERENCES

[1] F. Jia, S. Li, H. Zuo, and J. Shen, "Deep neural network ensemble for the intelligent fault diagnosis of machines under imbalanced data," *IEEE Access*, vol. 8, pp. 120974–120982, 2020.

[2] J. Shen, S. Li, F. Jia, H. Zuo, and J. Ma, "A deep multi-label learning framework for the intelligent fault diagnosis of machines," *IEEE Access*, vol. 8, pp. 113557–113566, 2020.

[3] X. Hu, H. Zhang, D. Ma, and R. Wang, "Status detection from spatial-temporal data in pipeline network using data transformation convolutional neural network," *Neurocomputing*, vol. 358, pp. 401–413, Sep. 2019.

[4] R. Wang, Q. Sun, D. Ma, D. Qin, Y. Gui, and P. Wang, "Line inductance stability operation domain assessment for weak grids with multiple constant power loads," *IEEE Trans. Energy Convers.*, early access, Sep. 2, 2020, doi: 10.1109/TEC.2020.3021070.

[5] Y. Wang, Z. Wei, and J. Yang, "Feature trend extraction and adaptive density peaks search for intelligent fault diagnosis of machines," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 105–115, Jan. 2019.

[6] Z. Wei, Y. Wang, S. He, and J. Bao, "A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection," *Knowl.-Based Syst.*, vol. 116, pp. 1–12, Jan. 2017.

[7] Z. Guo, Y. Wan, and H. Ye, "An unsupervised fault-detection method for railway turnouts," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 8881–8901, Nov. 2020.

[8] J. Dai, J. Wang, W. Huang, J. Shi, and Z. Zhu, "Machinery health monitoring based on unsupervised feature learning via generative adversarial networks," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 5, pp. 2252–2263, Oct. 2020.

[9] Q. Li, C. Shen, L. Chen, and Z. Zhu, "Knowledge mapping-based adversarial domain adaptation: A novel fault diagnosis method with high generalizability under variable working conditions," *Mech. Syst. Signal Process.*, vol. 147, Jan. 2021, Art. no. 107095.

[10] C. Hu, Y. Wang, and J. Gu, "Cross-domain intelligent fault classification of bearings based on tensor-aligned invariant subspace learning and two-dimensional convolutional neural networks," *Knowl.-Based Syst.*, vol. 209, Dec. 2020, Art. no. 106214.

[11] Z. Chen, G. He, J. Li, Y. Liao, K. Gryllias, and W. Li, "Domain adversarial transfer network for cross-domain fault diagnosis of rotary machinery," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 8702–8712, Nov. 2020.

[12] K. Yu, H. Han, Q. Fu, H. Ma, and J. Zeng, "Symmetric co-training based unsupervised domain adaptation approach for intelligent fault diagnosis of rolling bearing," *Meas. Sci. Technol.*, vol. 31, no. 11, Nov. 2020, Art. no. 115008.

[13] K. Yu, T. R. Lin, H. Ma, X. Li, and X. Li, "A multi-stage semi-supervised learning approach for intelligent fault diagnosis of rolling bearing using data augmentation and metric learning," *Mech. Syst. Signal Process.*, vol. 146, Jan. 2021, Art. no. 107043.

[14] T. Ko and H. Kim, "Fault classification in high-dimensional complex processes using semi-supervised deep convolutional generative models," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2868–2877, Apr. 2020.

[15] X. Tao, C. Ren, Q. Li, W. Guo, R. Liu, Q. He, and J. Zou, "Bearing defect diagnosis based on semi-supervised kernel local Fisher discriminant analysis using pseudo labels sciencedirect," *ISA Trans.*, Oct. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0019057820304365

[16] K. Yu, H. Ma, T. R. Lin, and X. Li, "A consistency regularization based semi-supervised learning approach for intelligent fault diagnosis of rolling bearing," *Measurement*, vol. 165, Dec. 2020, Art. no. 107987.

[17] Y. Zhang, Z. Ren, and S. Zhou, "An intelligent fault diagnosis for rolling bearing based on adversarial semi-supervised method," *IEEE Access*, vol. 8, pp. 149868–149877, 2020.

[18] Y. Xie, "Modified label propagation on manifold with applications to fault classification," *IEEE Access*, vol. 8, pp. 97771–97782, 2020.

[19] F. Morbidi, "The deformed consensus protocol," *Automatica*, vol. 49, no. 10, pp. 3049–3055, 2013.

[20] X. Zhu, Z. Ghahramani, and J. D. Lafferty, "Semi-supervised learning using Gaussian fields and harmonic functions," in *Proc. ICML*, 2003, pp. 912–919.

[21] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples," *J. Mach. Learn. Res.*, vol. 7, pp. 2399–2434, Nov. 2006.

[22] C. Gong, T. Liu, D. Tao, K. Fu, E. Tu, and J. Yang, "Deformed graph Laplacian for semisupervised learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2261–2274, Oct. 2015.

[23] H. Xu and S. Mannor, "Robustness and generalization," *Mach. Learn.*, vol. 86, no. 3, pp. 391–423, Mar. 2012.

[24] D. Colquhoun and A. G. Hawkes, "A q-matrix cookbook," in *Single-Channel Recording*. Boston, MA, USA: Springer, 1995, pp. 589–633.

[25] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16, 2004, pp. 321–328.

[26] H. Zhang, Z. Zhang, M. Zhao, Q. Ye, M. Zhang, and M. Wang, "Robust triple-matrix-recovery-based auto-weighted label propagation for classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4538–4552, Nov. 2020.

[27] J. Meng, D. Fu, and Y. Tang, "Belief-peaks clustering based on fuzzy label propagation," *Int. J. Speech Technol.*, vol. 50, no. 4, pp. 1259–1271, Apr. 2020.

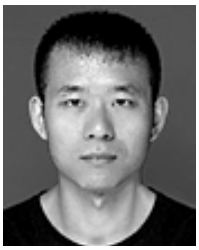
[28] Y. Zhang, Y. Liu, R. Jin, J. Tao, L. Chen, and X. Wu, "GLLPA: A graph layout based label propagation algorithm for community detection," *Knowl.-Based Syst.*, vol. 206, Oct. 2020, Art. no. 106363.

[29] W. Zhang, C. Yu, X. Wang, and F. Liu, "Predicting circrna-disease associations through linear neighborhood label propagation method," *IEEE Access*, vol. 7, pp. 83474–83483, 2019.

[30] L. Bai, J. Wang, J. Liang, and H. Du, "New label propagation algorithm with pairwise constraints," *Pattern Recognit.*, vol. 106, Oct. 2020, Art. no. 107411.
[31] C. Hu and Y. Wang, "Multidimensional denoising of rotating machine based on tensor factorization," *Mech. Syst. Signal Process.*, vol. 122, pp. 273–289, May 2019.
[32] C. Hu, S. He, and Y. Wang, "A classification method to detect faults in a rotating machinery based on kernelled support tensor machine and multilinear principal component analysis," *Int. J. Speech Technol.*, vol. 7, pp. 1–13, Nov. 2020.

JIAWEI GU received the master's degree in mechanical engineering from the Guilin University of Electronic Technology, Guilin, China, in 2020. His research interests include machine learning, fault diagnosis, transfer learning, and computer vision.

YANXUE WANG (Member, IEEE) received the Ph.D. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2009. From 2010 to 2011, he was a Postdoctoral Researcher with the University of Ottawa, Ottawa, ON, Canada. From 2013 to 2015, he was an Alexander von Humboldt Research Fellow with the Department of Strukturdynamik, Technische University at Darmstadt, Darmstadt, Germany. He is currently a Full Professor with the School of Mechanical Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing, China. He has authored/coauthored more than 50 technical papers published in prestigious international journals and conferences. His research interests include industrial big data analytics, mechanical fault diagnosis, and prognostics and health management (PHM).



CHAOFAN HU (Member, IEEE) received the Ph.D. degree in mechanical engineering from the Guilin University of Electronic Technology, Guilin, China, in 2020. He is currently a Full Associate Professor with the School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology. He has authored/coauthored more than 15 technical papers published in prestigious international journals and conferences. His research interests include industrial mechanical fault diagnosis and rensor representation.

ZEXI LUO received the B.S. degree in solid mechanics from Tongji University, Shanghai, China, in 2020. He is currently working as an Engineer with the Shanghai Institute of Aeronautical Measurement and Control Technology. His research interests include fault diagnosis and mechanical dynamic modeling.