sensors

MDPI

Article

A Semi-Supervised Adaptive Matrix Machine Approach for Fault Diagnosis in Railway Switch Machine

Wenqing Li, Zhongwei Xu *, Meng Mei, Meng Lan, Chuanzhen Liu and Xiao Gao

School of Electronic and Information Engineering, Tongji University, Shanghai 201804, China; 2310877@tongji.edu.cn (W.L.); mei_meng@163.com (M.M.); 2310912@tongji.edu.cn (M.L.); liuchuanzhen88@163.com (C.L.); xiao_gao1@163.com (X.G.)
* Correspondence: xuzhongweish@163.com

Abstract: The switch machine, an essential element of railway infrastructure, is crucial in maintaining the safety of railway operations. Traditional methods for fault diagnosis are constrained by their dependence on extensive labeled datasets. Semi-supervised learning (SSL), although a promising solution to the scarcity of samples, faces challenges such as the imbalance of pseudo-labels and inadequate data representation. In response, this paper presents the Semi-Supervised Adaptive Matrix Machine (SAMM) model, designed for the fault diagnosis of switch machine. SAMM amalgamates semi-supervised learning with adaptive technologies, leveraging adaptive low-rank regularizer to discern the fundamental links between the rows and columns of matrix data and applying adaptive penalty items to correct imbalances across sample categories. This model methodically enlarges its labeled dataset using probabilistic outputs and semi-supervised, automatically adjusting parameters to accommodate diverse data distributions and structural nuances. The SAMM model's optimization process employs the alternating direction method of multipliers (ADMM) to identify solutions efficiently. Experimental evidence from a dataset containing current signals from switch machines indicates that SAMM outperforms existing baseline models, demonstrating its exceptional status diagnostic capabilities in situations where labeled samples are scarce. Consequently, SAMM offers an innovative and effective approach to semi-supervised classification tasks involving matrix data.

Keywords: fault diagnosis; semi-supervised learning; adaptive matrix machine; pseudo-labeling imbalance; switch machine

check for updates

Citation: Li, W.; Xu, Z.; Mei, M.; Lan, M.; Liu, C.; Gao, X. A Semi-Supervised Adaptive Matrix Machine Approach for Fault Diagnosis in Railway Switch Machine. *Sensors* **2024**, *24*, 4402. https://doi.org/10.3390/s24134402

Academic Editor: Andrea Cataldo

Received: 23 April 2024
Revised: 3 July 2024
Accepted: 5 July 2024
Published: 7 July 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The turnout represents one of the three primary railroad outdoor components, with its condition having a direct impact on the safety of shunting and train station traffic [1]. The switch machine locks the track in either the directional or the reverse direction, serving as the motion execution unit of the turnout, as shown in Figure 1. Currently, the primary maintenance approach for turnout equipment is a combination of "cycle repair" and "fault repair" [2]. Maintenance personnel obtain the action current and power data of the switch machine through the centralized signaling monitoring (CSM) system and analyze the working status of the turnouts based on their professional knowledge and experience, thereby aiding in turnout maintenance. However, the conversion signals of switch machine equipment exhibit non-stationary and non-linear characteristics. The numerous types of turnout faults with complex characteristics make fault detection and classification very challenging. This maintenance approach has several drawbacks, including long fault delays, low fault diagnosis accuracy, and high labor intensity. Additionally, it often leads to "under-maintenance" and "over-maintenance," highlighting the limitations of the current maintenance model. Fault diagnosis of the switch machine is crucial in providing a reliability guarantee for the entire life cycle of the interlocking, attracting extensive attention from experts and scholars in the field of condition repair [3].

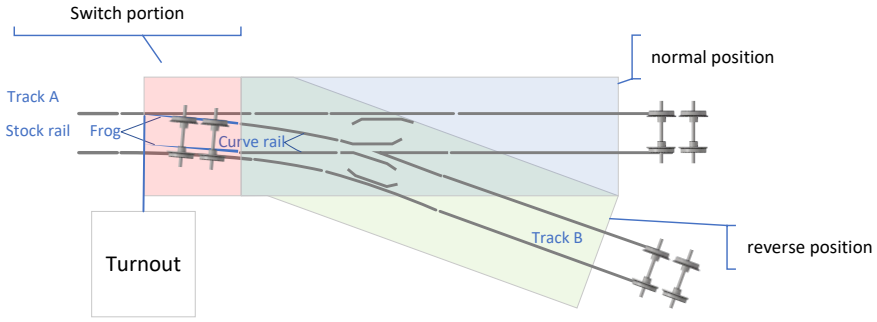


Figure 1. Turnout structural schematic.

With the advancement of intelligent operation and maintenance of railroad electric services, an increasing number of data-driven fault diagnosis methods have emerged for switch machines and other electromechanical equipment [4,5]. These methods are categorized into three types according to their feature classification strategy: (1) distance-based, determining the abnormality by setting standard curves [6,7]; (2) classifier-based, including support vector machine (SVM) [8], k-nearest neighbor(KNN), [9] and random forest (RF) [10]; (3) deep-learning-based methods, including convolutional auto-encoder (CAE) [11], convolutional neural networks (CNN) [12], and long short-term memory networks (LSTM) [13]. However, these methods demand substantial professional and field-specific expertise, and obtaining labeled data is both laborious and expensive, hindering intelligent fault diagnosis [14,15].

Since semi-supervised learning (SSL) offers a solution to requiring only a few labeled data, numerous studies have begun incorporating SSL to enhance fault diagnosis performance [16,17]. Lao et al. [18] introduced a semi-supervised weighted prototype network (SSWPN) targeting the issue of switch machine fault diagnosis with limited labeling. Shi et al. [19] extracted dynamic current profile features, integrating SVM with semi-supervised strategies to ascertain the turnout state. In semi-supervised learning, pseudo-labeling serves as an essential strategy. Initially, the model is trained using a limited dataset of labeled examples. Subsequently, it assigns labels to a substantial volume of unlabeled data based on predictions. Predictions made with high confidence are considered accurate labels and are then integrated into the further training of the model [20]. Although some studies have introduced semi-supervision into the field of fault diagnosis, semi-supervised fault diagnosis for switch machine troubleshooting remains nascent, facing several challenges. (1) Despite training with balanced data and evaluating balanced target data, an inherent imbalance in pseudo-labels emerges due to data similarity [21]. In classification models, the loss penalty plays a crucial role in defining the boundary between two hyperplanes. Regrettably, applying a single penalty parameter across all samples results in hyperplane shifts in unbalanced class distributions, with the model penalizing minority classes and favoring the majority class for hyperplane delineation. (2) Acquired signals are inherently represented as fault images, showcasing varied correlations among the matrix's rows and columns. However, training with solely vector features, reliant on expert knowledge, inevitably compromises the spatial feature integrity when vectorizing matrix samples.

Several scholars have embarked on studies to directly utilize matrix samples for modeling. Capturing the rank of a matrix, a measure of the correlation between its rows and columns, is crucial for constructing a matrix classifier. Various researchers have introduced diverse approaches to address the matrix rank issue, including decomposing the matrix into rank-k matrices [22,23] or constraining the matrix rank to 1 [24]. However, these techniques often necessitate predetermining the matrix rank, thus limiting their practicality. Conversely, support matrix machines (SMM) [25] suggest employing the kernel norm to approximate matrix ranks, enabling direct classification of 2D matrix features without presetting the ranks, thereby preserving the data's structural integrity. In recent years, efforts have been made to enhance SMM's performance through innovations like the multi-

class support matrix machine (MSMM) [26], multi-class support matrix machine based on evolutionary optimization (MSMM-CE) [27], security transfer support matrix machine (STSMM) [28], among others. Recent studies have demonstrated that approximating rankings with kernel norms is suboptimal [29,30]. The total of the singular values, each of which has a size indicative of its significance, forms the nuclear norm. Using the kernel norm to approximate matrix rank treats all singular values equally and lacks adaptiveness, significantly reducing its flexibility [31].

This study introduces a semi-supervised adaptive matrix machine (SAMM) method for diagnosing faults in switch machines, specifically applying it to analyze current signals from the ZDJ9 switch machine dataset. The diagnostic capabilities of this method have been validated through experimental comparisons, which demonstrate its superiority over other contemporary diagnostic approaches. The primary contributions of this research are threefold:

(1) The incorporation of an adaptive low-rank regularizer selectively retains larger singular values, improving the approximation of the matrix rank and enhancing the extraction of fundamental connections between the rows and columns of matrix data.
(2) The development of a probabilistic output strategy for SAMM, coupled with a semi-supervised learning (SSL) framework that utilizes these outputs to assign high-confidence pseudo-labels to unlabeled samples, effectively mitigating the challenges associated with a lack of labeled data.
(3) The introduction of an adaptive penalty term to address the imbalance in pseudo-label distribution, which adjusts the hinge loss penalty coefficient based on sample quantity to counteract learning biases.

The remainder of this paper is organized as follows: Section 2 offers a concise overview of the original SMM model. Section 3 describes the proposed SAMM model and its diagnostic framework in detail. Section 4 discusses the experimental validation of the method. Section 5 concludes the paper with a summary of the findings.

2. Support Matrix Machine

The support matrix machine (SMM) is a classification methodology specifically designed to handle input data in matrix forms, shown in Figure 2. Unlike conventional classification methods, which convert matrices into vectors and potentially compromise the structural integrity of the data, SMM preserves the matrix format. This retention enables SMM to fully utilize the structural information within matrices. It introduces a novel penalty term, the spectral elastic net, to leverage this advantage. By maintaining the matrix structure, SMM effectively captures the inherent structure and correlations of the data, thereby enhancing classification performance.

The objective function of SMM is formulated as follows:

$$\min_{W,b} \frac{1}{2} tr\left(\mathbf{W}^T\mathbf{W}\right) + \beta \parallel \mathbf{W}\parallel_* + \rho\sum_{i=1}^{N}\left\{1 - t_i\left[tr\left(\mathbf{W}^T\mathbf{A}_i\right) + b_i\right]\right\}_+ \quad (1)$$

where $\mathbf{A}_i \in \mathbb{R}^{p\times q}, i = 1, 2, \ldots, N$ represents the training matrix data. $y_i \in \{-1, 1\}$ denotes the corresponding labels. The regression matrices $\mathbf{W} \in \mathbb{R}^{p\times q}$ are inversely proportional to the distance of the hyperplane margin. b signifies the bias. β and ρ are two hyperparameters.

The objective function of the support matrix machine (SMM) comprises two principal components: the matrix-form hinge loss and the spectral elastic net penalty. Hinge loss, a common feature in classification models, promotes sparsity and robustness. In SMM, this loss function measures the classification error by calculating the discrepancy between the model's predictions and the actual labels for each training sample. The spectral elastic net penalty, integral to SMM, exploits the structural information of the feature matrix, capturing correlations within its columns and rows. In particular, $tr\left(\mathbf{W}^T\mathbf{W}\right)$ is employed to keep the model's complexity under control and avoid overfitting, ensuring that the model adheres to the rule of minimizing structural risk. The kernel norm $\parallel \mathbf{W}\parallel_*$, defined as the total of the singular value decompositions of the matrix, serves as a metric for assessing the

matrix's low rank, which facilitates the extraction of structural information from matrix data [32]. However, minimizing the kernel norm by substituting the rank function with an approximation may compromise accuracy, especially in matrices with complex structures. The kernel norm represents only a relaxed approximation of the rank function, and it does not fully capture the matrix's true rank. Significant deviations from the true rank occur within the kernel norm when the singular values differ from 1 [33]. This relaxation causes the kernel norm to over-penalize the matrix, potentially yielding a suboptimal low-rank approximation.

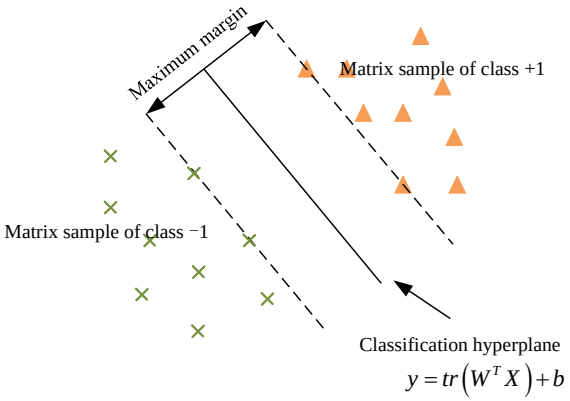


Figure 2. Classification principle of SMM.

3. Semi-Supervised Adaptive Matrix Machine

In this section, the proposed semi-supervised adaptive matrix machine (SAMM) method is presented, which aims to address the challenges of insufficient labeled samples and pseudo-label imbalance, as well as to more accurately capture the correlations between matrix rows and columns. To resolve the non-smooth optimization issue within the SAMM model, an alternating update strategy is employed [34], facilitating the efficient solution of the model and the attainment of the optimal solution.

3.1. SAMM Model

The semi-supervised adaptive matrix machine (SAMM) combines semi-supervised learning with adaptive techniques, utilizing an adaptive low-rank regularizer to identify correlation within matrix data and an adaptive penalty term to mitigate the impact of inter-class samples on hyperplane margins, as shown in Figure 3. By integrating probabilistic output, which incrementally expands the labeled sample set, SAMM not only gradually improves classification performance but also autonomously adjusts model parameters during the learning phase to suit various data distributions and feature architectures. The objective function for SAMM is delineated as follows:

$$\begin{aligned} &\min_{\mathbf{W},b} \frac{1}{2} tr\left(\mathbf{W}^T\mathbf{W}\right) + \beta \sum_{k=1}^{r} \log|\sigma_k + \varepsilon| \\ &+ \rho_i \sum_{i=1}^{N_l} \left\{1 - y_i\left[tr\left(\mathbf{W}^T\mathbf{A}_i\right) + b\right]\right\}_+ + \rho_j \sum_{j=1}^{Nr} \left\{1 - \tilde{y}_j\left[tr\left(\mathbf{W}^T\tilde{\mathbf{A}}_j\right) + b\right]\right\}_+ \end{aligned} \tag{2}$$

The first term influences the model in the same way as observed in the support matrix machine (SMM). The second term $\beta \sum_{k=1}^{r} \log|\sigma_k + \varepsilon|$ introduces an adaptive low-rank regularizer, where ε is a sufficiently small positive number, ensuring $\sigma_k + \varepsilon$ is not zero. Larger singular values correspond to row and column information within the matrix and should be preserved, whereas smaller singular values, which are often linked to irrelevant or redundant data, should be discarded. Adaptive low-rank regularizers maintain these larger singular values and reduce the smaller ones to zero or near-zero values. By minimizing

this regularizer, SAMM effectively extracts low-rank matrix information and adaptively selects and preserves singular values associated with highly correlated data. Utilizing the adaptive low-rank regularizer allows SAMM to more accurately estimate the matrix's rank and extract strong correlations between rows and columns from the matrix data. This adaptivity enables SAMM to handle matrix-form data more effectively and improve its classification performance.

The adaptive penalty term computes category weights according to the ratio of pseudo-labeled samples per category, adjusting the hinge loss's penalty parameter ρ_i based on these weights to optimize the handling of unbalanced samples. If samples are balanced, the penalty remains constant $\rho_i = \rho$; for unbalanced samples, ρ_i can be determined by

$$\rho_i = \begin{cases} \rho \frac{N_2}{N}, i \in N_1 \\ \rho(1 + \frac{N_1}{N}), i \in N_2 \end{cases} \tag{3}$$

$N1$ and $N2$ represent the number of samples in the majority and minority categories, respectively, while N denotes the combined total of samples from both categories. With the introduction of an adaptive penalty term, samples from the majority categories incur a lower penalty than those from the minority categories. Consequently, the SAMM model effectively considers the features of all categories within unbalanced datasets, thereby avoiding the problem of overemphasizing the majority categories while neglecting the minority ones.

To address the challenge of recognizing switch machine fault status with a limited number of labeled samples, we leverage both a small set of labeled and a substantial pool of unlabeled samples. We have developed a semi-supervised model that integrates probabilistic outputs, utilizing the SAMM model. Initially, the model undergoes training with the labeled dataset $\mathcal{A}_l = \{\mathbf{A}_i, y_i\}_{i=1}^{N_l}$. The Platt Scaling [35] method is utilized to map the output of the SAMM model for each sample into the [0,1] interval, serving as a probability estimate of the sample's category membership. Utilizing the Wu method [36], we couple $C(C-1)/2$ SAMM probability estimates pairwise into a single value, wherein the maximum probability output indicates the unlabeled sample's confidence level for true category membership. Confidence thresholding is a widely used technique to enhance pseudo-labeling. By setting a higher threshold, the reliability of pseudo-labels is improved [37]. A large unlabeled dataset $\mathcal{A}_u = \{\mathbf{A}_j\}_{j=1}^{N_u}$ is provided, from which samples that exceed a specified confidence threshold are incorporated into the labeled dataset as reliable samples $\mathcal{A}_r = \{\tilde{\mathbf{A}}_j, \tilde{y}_j\}_{j=1}^{N_r}$.Repeat this process until no samples exceed the confidence threshold or the maximum iteration count is achieved. This semi-supervised learning approach effectively addresses the switch machine fault diagnosis challenge with few labeled samples, diminishing the time and economic expenditures associated with sample labeling.

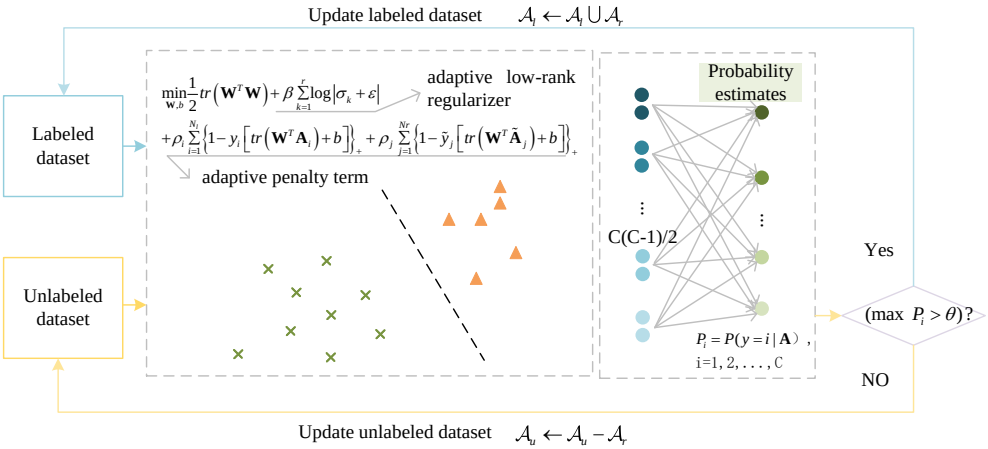


Figure 3. SAMM model.

3.2. SAMM Learning Algorithm

Solving the SAMM model presents a non-smooth optimization challenge, complicating the search for a globally optimal solution. To address this issue, the alternating direction multiplier method (ADMM) is introduced as an effective algorithm for the SAMM model's resolution. The ADMM algorithm tackles the original challenge by breaking it down into two subproblems and applying iterative alternating updates. During each iteration, the ADMM algorithm incrementally approaches the optimal solution by updating primal and dual variables. This iterative procedure efficiently resolves the SAMM model, achieving an optimal solution characterized by low rank and sparsity. In the ADMM framework, Equation (2) is reformulated as Equation (4).

$$\begin{array}{cc} \arg\min\limits_{\mathbf{W},b,\mathbf{Z}} & h(\mathbf{W},b)+g(\mathbf{Z}) \\ \text{s.t.} & \mathbf{Z}-\mathbf{W}=0 \end{array} \tag{4}$$

here

$h(\mathbf{W},b)=\frac{1}{2}tr\left(\mathbf{W}^T\mathbf{W}\right)+\rho_i\sum\limits_{i=1}^{N_l}\left\{1-y_i\left[tr\left(\mathbf{W}^T\mathbf{A}_i\right)+b\right]\right\}_+ +\rho_j\sum\limits_{j=1}^{Nr}\left\{1-\tilde{y}_j\left[tr\left(\mathbf{W}^T\tilde{\mathbf{A}}_j\right)+b\right]\right\}_+$

$g(\mathbf{Z})=\tau\sum\limits_{k=1}^{r}\log|\sigma_k+\varepsilon|$ σ_k is the kth singular value of the matrix $\mathbf{Z}$, $k=1,2,...,r$

The augmented Lagrangian function is subsequently defined as follows

$$L(\mathbf{W},b,\mathbf{Z},\mathbf{\Lambda})=h(\mathbf{W},b)+g(\mathbf{Z})+\langle\mathbf{\Lambda},\mathbf{Z}-\mathbf{W}\rangle+\frac{\delta}{2}\parallel\mathbf{Z}-\mathbf{W}\parallel_F^2 \tag{5}$$

$\delta>0$ represents the step size, and $\mathbf{\Lambda}$ denotes the Lagrange multiplier. Following the ADMM framework, the objective function is divided into two subproblems (concerning $\mathbf{Z}$ and $\mathbf{W}$) and resolved through iterative computation. During each iteration, the solver sequentially minimizes $\mathbf{Z}$ and $\mathbf{W}$, followed by an update to the Lagrange multipliers in alignment with these adjustments. $\mathbf{W}$, $\mathbf{Z}$, and $\mathbf{\Lambda}$ are updated as follows.

$$\begin{cases} \mathbf{Z}^{(t+1)}=\arg\min L\left(\mathbf{Z},\mathbf{W}^{(t)},\mathbf{\Lambda}^{(t)}\right) \\ \mathbf{W}^{(t+1)}=\arg\min L\left(\mathbf{Z}^{(t+1)},\mathbf{W},\mathbf{\Lambda}^{(t)}\right) \\ \mathbf{\Lambda}^{(t+1)}=\mathbf{\Lambda}^{(t)}+\delta\left(\mathbf{Z}^{(t+1)}-\mathbf{W}^{(t+1)}\right) \end{cases} \tag{6}$$

Here t and t + 1 signify the tth and t + 1th iterations, respectively.

(1) To solve the subproblem of $\mathbf{Z}$, assume $(\mathbf{W},b)$ and $\mathbf{\Lambda}$ are held constant, reducing it to a function concerning $\mathbf{Z}$ expressed as:

$$\begin{aligned} L_{\mathbf{Z}}=&g(\mathbf{Z})+\langle\mathbf{\Lambda},\mathbf{Z}-\mathbf{W}\rangle+\frac{\delta}{2}\parallel\mathbf{Z}-\mathbf{W}\parallel_F^2 \\ =&g(\mathbf{Z})+\frac{\delta}{2}\left\|\mathbf{Z}-\mathbf{W}+\frac{\mathbf{\Lambda}}{\delta}\right\|_F^2 \end{aligned} \tag{7}$$

Let $\mathbf{I}=\mathbf{W}-\frac{\mathbf{\Lambda}}{\delta}$ undergo singular value decomposition (SVD) in the following manner: $\mathbf{I}=\mathbf{U_I}\mathbf{\Sigma_I}\mathbf{V_I^T}$. According to [31], $\mathbf{Z}$ can be solved as

$$\mathbf{Z}^*=\mathbf{U_I}Prox_\tau\left(\mathbf{\Sigma}_{\mathbf{I}_k}\right)\mathbf{V}_\mathbf{I}^T \tag{8}$$

where the nearest neighbor operator

$$Prox_\tau(\sigma_i)=\begin{cases}\max\left(\sigma_i-\frac{2\tau}{\sigma_i+\sqrt{\sigma_i^2-4\tau}},0\right), i\in I_1 \\ 0, i\in I_2\end{cases}$$

here $I_1=\left\{i\mid\sigma_i^2-4\tau\geq0\right\}$, $I_2=\left\{i\mid\sigma_i^2-4\tau<0\right\}$

(2) To address the subproblem concerning **W**, we undertake minimization of the expression encapsulating all terms associated with **W** as outlined in Equation (5).

$$L_{\mathbf{W}} = h(\mathbf{W}, b) + \langle \mathbf{\Lambda}, \mathbf{Z} - \mathbf{W} \rangle + \frac{\delta}{2} \parallel \mathbf{Z} - \mathbf{W} \parallel_F^2 \tag{9}$$

assume

$$\begin{gathered} \min_{\mathbf{W},b} \tfrac{1}{2} tr(\mathbf{W}^T\mathbf{W}) + \sum_{i=1}^{N} \rho_i \xi_i + \langle \mathbf{\Lambda}, \mathbf{Z} - \mathbf{W} \rangle + \tfrac{\delta}{2} \parallel \mathbf{Z} - \mathbf{W} \parallel_F^2 \\ s.t. y_i \left[tr(\mathbf{W}^{\mathrm{T}} \mathbf{A}_i) + b \right] \geq 1 - \xi_i \\ \xi_i \geq 0, i = 1, 2, \ldots, N \end{gathered} \tag{10}$$

Constructed via the Lagrange multiplier method

$$\begin{gathered} L(\mathbf{W}, b, \xi, \alpha, \beta) = \tfrac{1}{2} tr(\mathbf{W}^T\mathbf{W}) + \sum_{i=1}^{N} \rho_i \xi_i + \langle \mathbf{\Lambda}, \mathbf{Z} - \mathbf{W} \rangle + \tfrac{\delta}{2} \parallel \mathbf{Z} - \mathbf{W} \parallel_F^2 \\ - \sum_{i=1}^{N} \alpha_i \left\{ y_i \left[tr(\mathbf{W}^{\mathrm{T}} \mathbf{A}_i) + b \right] - 1 + \xi_i \right\} - \sum_{i=1}^{N} \beta_i \xi_i \end{gathered} \tag{11}$$

with partial derivatives set to zero for b and β_i.

$$\begin{cases} \dfrac{\partial L}{\partial b} = \displaystyle\sum_{i=1}^{N} \alpha_i y_i = 0 \\ \dfrac{\partial L}{\partial \xi_i} = \displaystyle\sum_{i=1}^{N} \rho_i - \sum_{i=1}^{N} \beta_i = 0 \end{cases} \tag{12}$$

Upon substituting Equation (12) into Equation (11), we obtain:

$$\begin{gathered} L(\mathbf{W}, b, \xi, \alpha, \beta) = \tfrac{1}{2} tr(\mathbf{W}^T\mathbf{W}) + \langle \mathbf{\Lambda}, \mathbf{Z} - \mathbf{W} \rangle + \tfrac{\delta}{2} \parallel \mathbf{Z} - \mathbf{W} \parallel_F^2 \\ - \sum_{i=1}^{N} \alpha_i \left\{ y_i \left[tr(\mathbf{W}^{\mathrm{T}} \mathbf{A}_i) + b \right] - 1 \right\} \end{gathered} \tag{13}$$

Equation (13) results from differentiating concerning **W** and setting the derivative to zero, yielding:

$$\mathbf{W}^* = \frac{1}{\delta + 1} \left(\mathbf{\Lambda} + \delta \mathbf{Z} + \sum_{i=1}^{N} \alpha_i y_i \mathbf{A}_i \right) \tag{14}$$

By reinserting Equations (12) and (14) into Equation (10), we derive the optimization problem for α as follows:

$$\begin{gathered} \max_{\alpha} \quad -\tfrac{1}{2} \alpha^{\mathrm{T}} \mathbf{R} \alpha + \mathbf{r}^{\mathrm{T}} \alpha \\ \text{s.t.} \quad \sum_{i=1}^{N} \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq \rho_i, i = 1, 2, \ldots, N \end{gathered} \tag{15}$$

Here $\mathbf{R} = \left[R_{ij} \right] \in \mathbb{R}^{p \times q}$, $\mathbf{r} = [r_i] \in \mathbb{R}^p$

$$\begin{cases} R_{ij} = \frac{y_i y_j tr\left(\mathbf{A}_i^{\mathrm{T}} \mathbf{A}_j \right)}{\delta + 1} \\ r_i = 1 - \frac{y_i tr\left[(\mathbf{\Lambda} + \delta \mathbf{Z})^{\mathrm{T}} \mathbf{A}_i \right]}{\delta + 1} \end{cases} \tag{16}$$

The optimal value for b is determined by defining an average solution as specified in

$$b^* = \frac{1}{|\mathbf{Z}^*|} \sum_{i \in \mathbf{S}^*} \left\{ y_i - tr\left[(\mathbf{W}^*)^{\mathrm{T}} \mathbf{A}_i \right] \right\} \tag{17}$$

here $\mathbf{Z}^* = \{ i : 0 < \alpha_i < C \}$

Algorithm 1 outlines the proposed learning algorithm for SAMM.

Algorithm 1: The learning algorithm for SAMM

Input: Training set $\{\mathbf{A}_i, y_i\}_{i=1}^{N}$, low-rank coefficient β, loss penalty coefficient ρ, step size δ.
Output $\mathbf{W}$, b
1. Initialize. $\mathbf{W}^{(0)} = \mathbf{0}, \mathbf{Z}^{(0)} = \mathbf{0}, \Lambda^{(0)} = \mathbf{0}, k = 1$
While not converging do
2. Update $\mathbf{Z}^k$ with Equation (8)
3. Update $\mathbf{W}^k$ and b with Equations (14) and (17)
4. Update $\mathbf{\Lambda}^{(k)}$ with Equation (6)
5. $k = k + 1$
End
6. **Return** $\mathbf{W}$,b

3.3. Fault Diagnosis Framework

The comprehensive framework of the model proposed herein is depicted in Figure 4, with the principal steps summarized as follows:

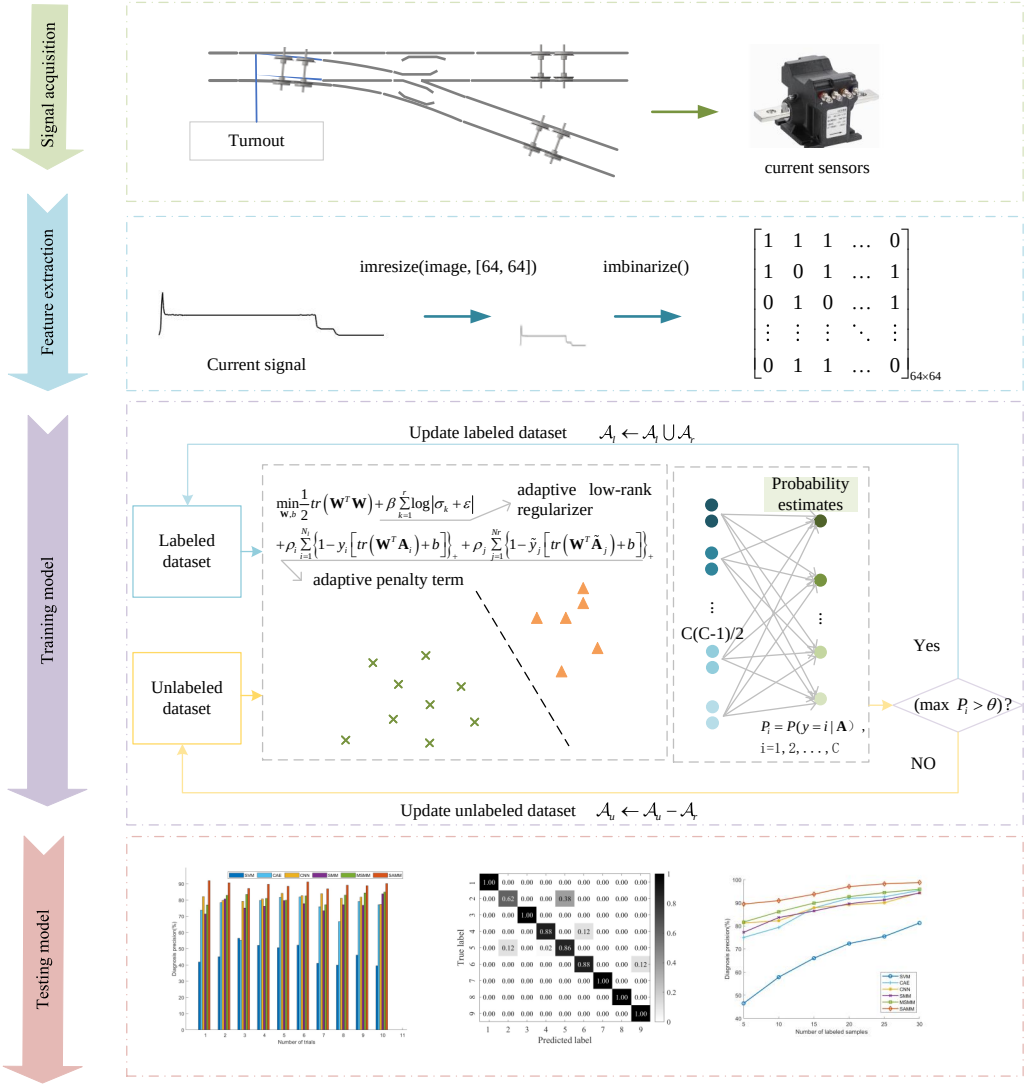


Figure 4. Entire framework of the proposed fault diagnosis approach.

Step 1: Signal acquisition. Acquire current signals of the switch machine across various fault states.

Step 2: Feature extraction. Convert continuous current signals into 2D matrix samples via downsampling and binarization techniques, enabling efficient processing and model training.
Step 3: Train the SAMM Model. Labeled and unlabeled samples from the training dataset are used to build the SAMM model. The model integrates an adaptive low-rank regularizer with an adaptive penalty term, enhancing matrix structure information extraction, and addressing the pseudo-labeling imbalance challenge of semi-supervised learning.
Step 4: Test the SAMM Model. Predict the switch machine's fault status by inputting test samples into the SAMM model.

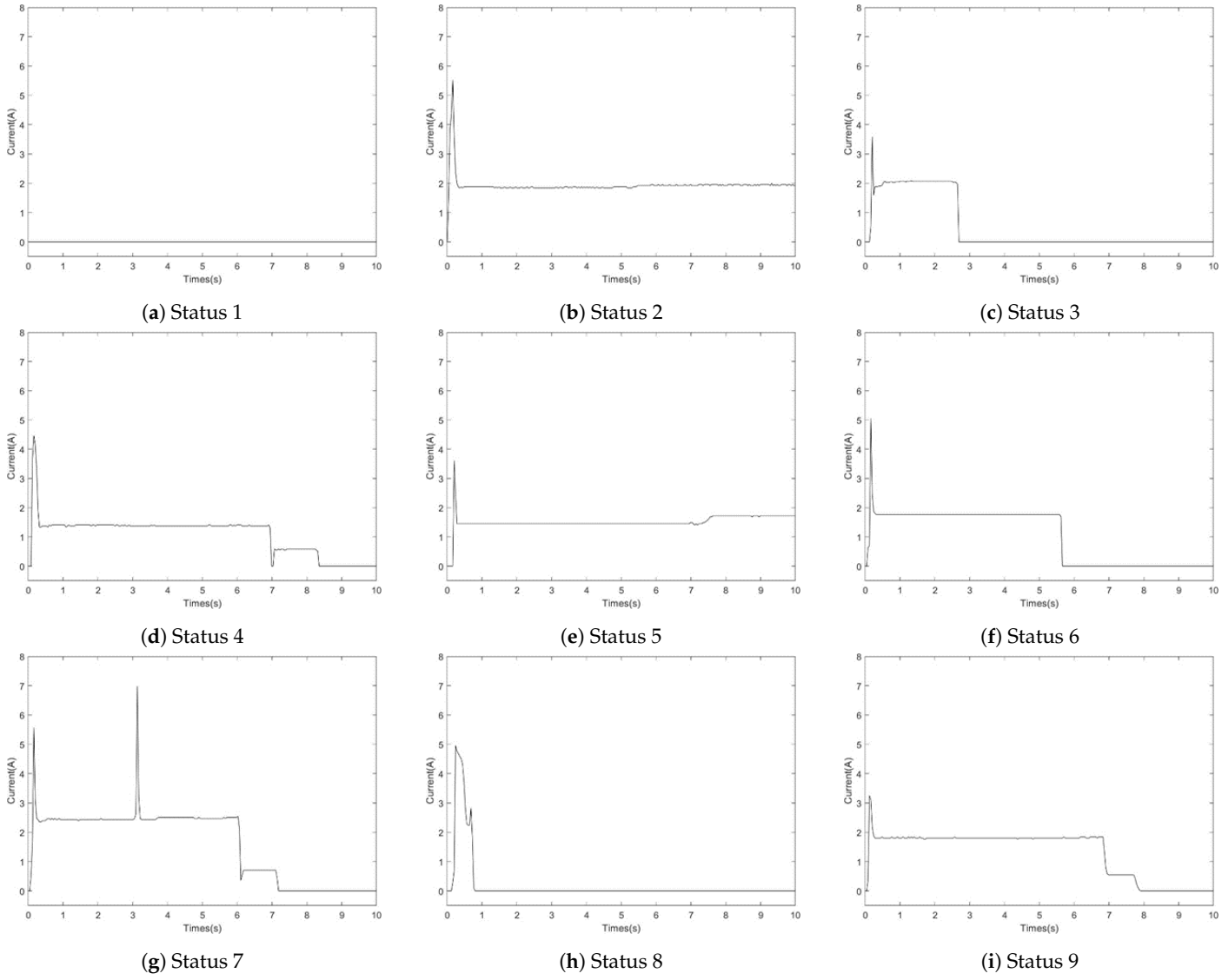
4. Experimental and Discussion

4.1. Description of the Data Set

The dataset originates from current signals generated by ZDJ9-type switch machines within the urban subway system. The SAMM model leverages current signals for fault diagnosis due to their direct correlation with the operational status of the railway switch machine. Although vibration and sound signals are also used in switch machine fault diagnosis [38,39], they present challenges in data collection and interpretation due to environmental noise and the need for precise sensor placement. Current signals, on the other hand, can be obtained through the CSM system, ensuring they are readily available and less susceptible to external noise. This method ensures minimal disruption to the switch machine's operation. This study's dataset was compiled by CASCO, a professional rail transit control system integrator, at specific stations along Shanghai Metro Line 13. The ZDJ9 switch machine uses a 380 V three-phase AC power supply, with phase currents A, B, and C supplying essential electrical power. This model completes a full state change in approximately 7 to 9 s, with current signals sampled at 25 Hz throughout this duration. A typical current profile encompasses four principal phases: unlocking, transition, locking, and slow release. For this study, the A-phase current curve was selected for dataset construction due to its comprehensive representation of the switch machine's motion. As detailed in Table 1 and Figure 5, the A-phase current dataset spans nine distinct fault statuses, comprising eight fault states and one normal state. Throughout the experiment, labeled training samples per fault status varied from 5 to 30, unlabeled training samples from 45 to 25, with a constant 50 test samples. Employing down sampling and binarization techniques, each raw current signal image was transformed into a 64 × 64-dimensional feature matrix, facilitating further processing and model training.

Table 1. Fault status phenomena and causes of ZDJ9 switch machine.

Label	Fault Phenomenon	Fault Cause
1	Consistently no current	Action circuit malfunction
2	The current remains constant during release	Mechanical resistance encountered
3	A sudden drop in current to zero	Insufficient contact or unlocked
4	The current release time is delayed	Abnormal motor condition
5	Increase in current during release.	Internal jamming and friction increase
6	Release without small steps	Abnormality in the indicated circuit
7	Pulse observed during switching	Poor contact of automatic switch
8	The curve only maintains 0∼1 s	Phase failure in the starting circuit
9	Normal state	Normal



(**a**) Status 1 (**b**) Status 2 (**c**) Status 3

(**d**) Status 4 (**e**) Status 5 (**f**) Status 6

(**g**) Status 7 (**h**) Status 8 (**i**) Status 9

Figure 5. Fault status current curves of ZDJ9 switch machine.

4.2. Comparison Experiment

To optimize the classification performance of the SAMM model, three key parameters were precisely adjusted in the experiments: the low-rank coefficient β, the loss penalty coefficient ρ, and the step size δ. In the experiment, δ was set to 0.01. A 5-fold cross-validation method was utilized to select the structural parameters β and ρ from the set $\{2^{-5}, 2^{-4.5}, \cdots, 2^{5}\}$, and the confidence threshold θ was set within the range of $\{0.5, 0.55, \cdots, 0.95\}$. To guarantee fairness and comparability in our experimental outcomes, structural parameters for each model were optimized before undertaking fault diagnosis tasks, ensuring optimal operation across differing models. The identical parameter optimization process was applied to other comparative models, notably support vector machines (SVM), support matrix machines (SMM), and multi-class support matrix machines (MSMM). The models' optimal parameters were determined using a 5-fold cross-validation technique, and the structural parameters' value ranges were determined by consulting relevant literature. Structural parameters for deep learning models such as the convolutional auto-encoder (CAE) and convolutional neural network (CNN) were selected based on insights gleaned from relevant literature. All diagnostic models operated within a Windows 11 (64-bit) and Matlab 2023a software environment. The utilized PC's hardware configuration primarily included an Intel(R) Core(TM) i7-13700H CPU and 32.0 GB RAM.

To thoroughly assess the classification performance across various classifiers, three evaluation metrics were employed: precision rate, recall rate, and F1 score. The precision

rate quantifies the proportion of accurately identified positive class samples among those deemed to be in a positive class, reflecting the classifier's accuracy. Conversely, recall gauges the proportion of all correctly identified samples within the actually positive class, indicating the classifier's coverage. The F1 score, a harmonized mean of precision and recall rates, serves as a singular comprehensive metric for gauging the classifier's overall effectiveness. Precision, recall, and F1 score, metrics suited for binary classification, were computed for each category using a macro-averaging approach and then averaged. For multicategory classification, these metrics are generalized from those utilized for k-category classification, as delineated in Table 2, and are defined as follows.

$$\begin{aligned} \text{precision} &= \frac{1}{k}\sum_{c=1}^{k}\frac{tp_c}{tp_c + fp_c}, \\ \text{recall} &= \frac{1}{k}\sum_{c=1}^{k}\frac{tp_c}{tp_c + fn_c}, \\ F_1 &= \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \end{aligned} \tag{18}$$

where tp_c, fp_c, fn_c, tn_c are true positives, false positives, false negatives, and true negatives within category c.

Table 2. Binary classification confusion matrix.

Data Class	Classified as Pos.	Classified as Neg.
pos	true positive (tp)	false negative (fn)
neg	false positive (fp)	true negative (tn)

To guarantee the results' reliability, each method was replicated 10 times for every sample case. Repeating the experiments aids in mitigating bias from random factors, thereby enhancing the robustness and credibility of the outcomes. Figure 6 displays the fault diagnosis precision for each model with merely five labeled training samples, showcasing that the SAMM model consistently outperforms others in terms of diagnostic precision across all experiments. The results demonstrate that the SAMM model sustains high diagnostic precision, even with a scarce quantity of labeled samples. Comparative results between SVM and matrix learning models (SMM, MSMM) illustrate that leveraging the structural information of images indeed enhances fault diagnosis performance. Given that SAMM adaptively leverages image structural information and mitigates the challenge of insufficient labeled samples, its overall diagnostic efficacy significantly surpasses that of the comparative models.

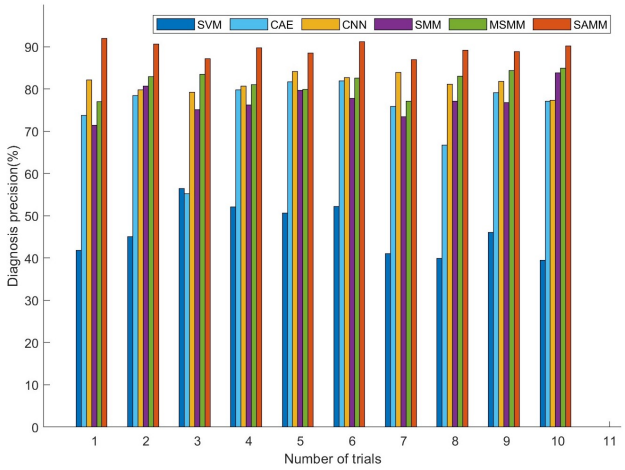
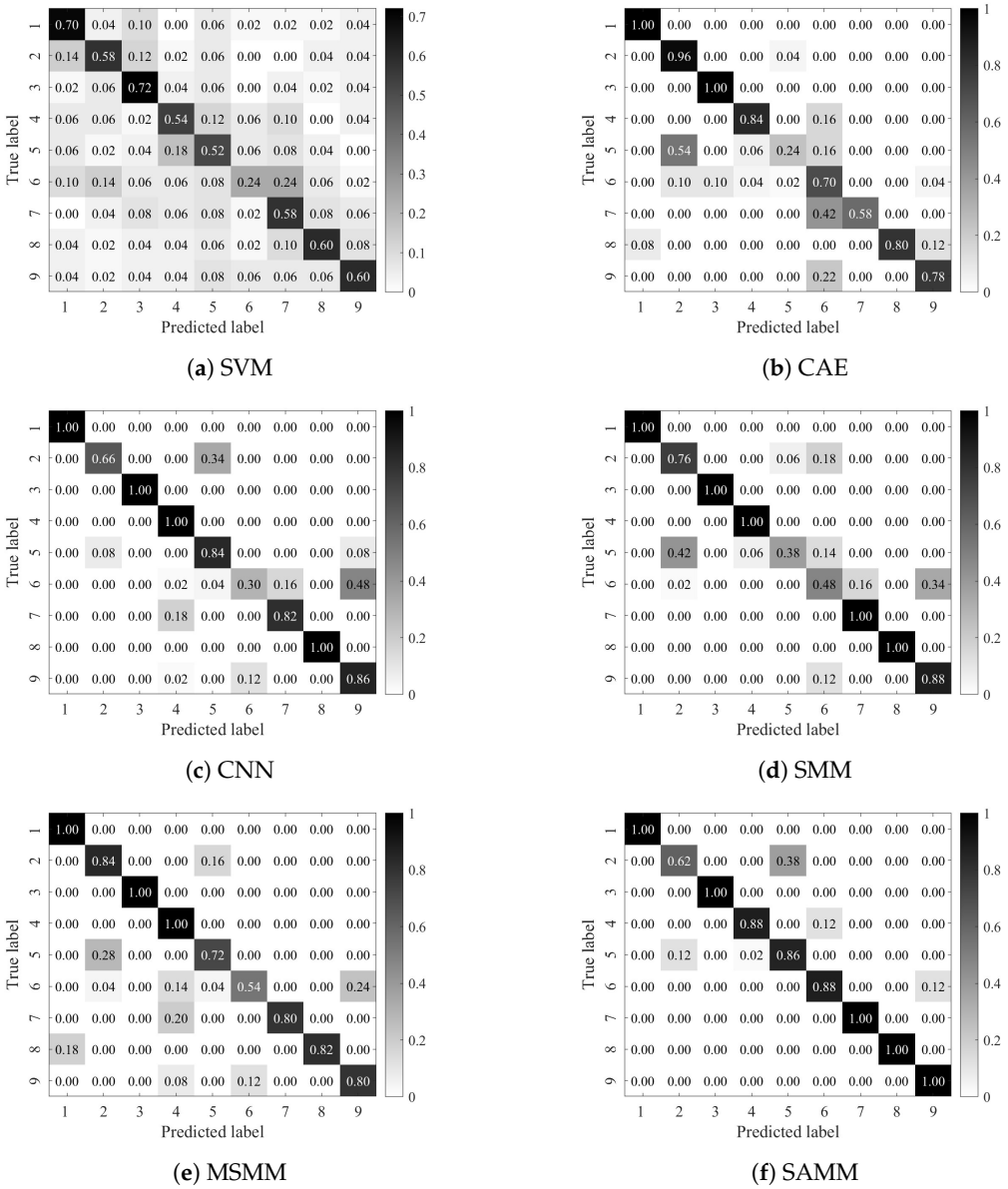


Figure 6. Repeat 10 times with 5 labels.

Figure 7 presents the confusion matrix for the optimal diagnostic outcomes across each model. The confusion matrix reveals that the SAMM model excels in identifying various fault statuses. The highest diagnostic accuracies achieved by SVM, CAE, CNN, SMM, MSMM, and SAMM are 56.45%, 82.00%, 84.15%, 83.75%, 85.00%, and 92.02%, respectively. The traditional SVM model's diagnostic accuracy significantly trails behind other models due to its inability to fully leverage image data's structural information. Despite the deep-learning-based CAE and CNN models' capability to extract higher-order image features, the scarcity of labeled samples limits their accuracy from reaching the desired level. By harnessing the structural features of image data, the matrix learning models SMM and MSMM outperform CAE and CNN, albeit with certain limitations. Conversely, the SAMM model adeptly extracts low-rank structural information from matrix samples and addresses category imbalance with adaptive penalty terms, achieving a leading diagnostic accuracy of 92.02%. It showcases superior fault diagnosis performance, even with a limited number of labeled samples. This underscores the SAMM model's advantages and efficacy in recognizing switch machine status.



(**a**) SVM (**b**) CAE

(**c**) CNN (**d**) SMM

(**e**) MSMM (**f**) SAMM

Figure 7. Confusion matrix of the optimal results for each model.

Figure 8 illustrates the fault diagnosis accuracy for each model across varying counts of labeled training samples. The figure demonstrates that the diagnostic accuracy for all models improves to varying extents with an increase in labeled training samples, aligning with the inherent reliance of machine learning models on the volume of training data. Significantly, the SAMM model's diagnostic accuracy surpasses that of the comparative models in every instance. Notably, across 5, 10, 15, 20, 25, and 30 labeled samples per fault status, the SAMM model achieved average diagnostic accuracies of 89.47%, 90.96%, 93.71%, 97.04%, 98.23%, and 98.80%, respectively, outperforming the lower accuracies recorded by the other models. The SAMM model's exceptional diagnostic performance is credited to its utilization of an ADMM-based solver, facilitating stable convergence to the global optimum and maximizing the model's potential. Crucially, SAMM's integration of an adaptive low-rank regularizer with an adaptive penalty term enables precise extraction of intrinsic low-rank structural information from matrix data and effectively addresses the prevalent issue of category imbalance in semi-supervised learning. Experimental findings indicate that SAMM's adaptive semi-supervised learning approach is particularly effective with a limited number of labeled samples. Remarkably, even with as few as five labeled samples, SAMM achieves a diagnostic accuracy of 89.47%, whereas other models exhibit a significant decline in performance. This affirms SAMM's superiority and practical utility in addressing the challenge of scarce labeled samples.

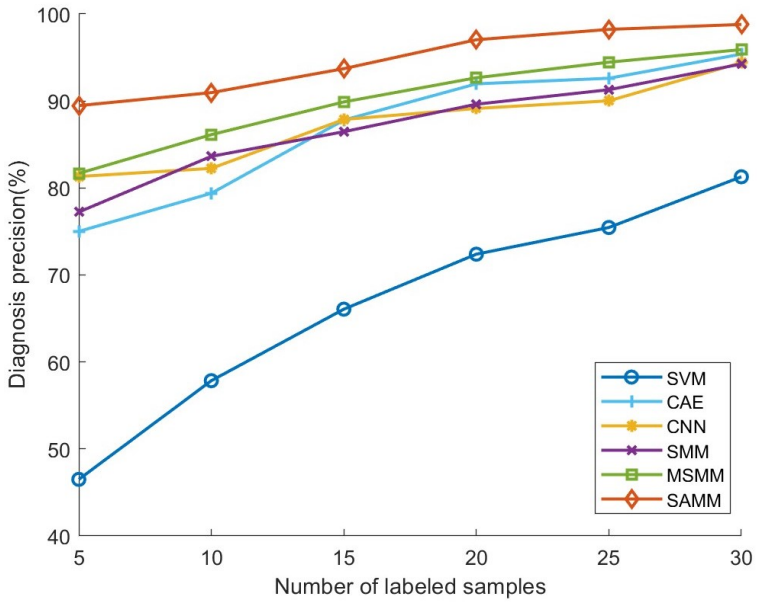


Figure 8. Fault diagnosis precision under different labeled samples.

Tables 3 and 4 detail the recall and F1 scores, respectively, for each model across varying numbers of labeled samples. When combined with the precision outcomes previously analyzed (Figure 8), a comprehensive evaluation of the models' overall diagnostic performance is facilitated. The recall and F1 score outcomes reveal that the SAMM model consistently outperforms all comparison models across various counts of labeled samples. With an increase in the number of labeled samples, while the recall and F1 scores for all models improve, the SAMM model's lead persists. Integrating the experimental findings of precision, recall, and F1 score, it becomes evident that the SAMM model's diagnostic efficacy surpasses that of other comparative models under scenarios with a limited number of labeled samples. This suggests that the strategies of employing an adaptive low-rank regularizer and adaptive penalty term enable the SAMM model to effectively tackle the challenges of scarce labeled samples and category imbalance, thereby demonstrating robust semi-supervised learning (SSL) capabilities.

Table 3. Recall rate of Different Models with Varying Number of Labeled Samples.

Model	Number of Labeled Samples in Each Status					
	5	**10**	**15**	**20**	**25**	**30**
SVM	40.69%	54.67%	65.11%	68.22%	71.78%	80.44%
CAE	72.71%	77.33%	87.11%	91.78%	91.56%	95.33%
CNN	79.25%	80.89%	85.78%	88.00%	88.00%	94.22%
SMM	74.89%	82.22%	84.67%	89.33%	89.33%	93.56%
MSMM	79.78%	82.00%	83.78%	92.44%	94.44%	95.11%
SAMM	87.91%	90.44%	93.56%	96.89%	98.22%	98.56%

Table 4. F1 score of different models with varying numbers of labeled samples.

Model	Number of Labeled Samples in Each Status					
	5	**10**	**15**	**20**	**25**	**30**
SVM	43.25%	56.21%	65.59%	70.25%	73.57%	80.87%
CAE	73.83%	78.35%	87.46%	91.88%	92.08%	95.36%
CNN	80.24%	81.57%	86.81%	88.57%	89.00%	94.32%
SMM	76.04%	82.92%	85.56%	89.48%	90.31%	93.90%
MSMM	80.57%	84.01%	86.73%	92.56%	94.45%	95.51%
SAMM	89.47%	90.70%	93.64%	96.96%	98.23%	98.68%

The experimental outcomes comprehensively illustrate that the SAMM model optimally utilizes the intrinsic structural information of image data. Concurrently, it addresses the challenges of scarce labeled samples and category imbalance through a semi-supervised learning strategy and adaptive mechanisms, culminating in superior performance in switch machine fault diagnosis compared to other models.Despite measurement noise and interference in real-world conditions, our method has demonstrated excellent fault diagnosis capabilities in experimental validations. The results indicate that, even with some noise and interference, the SAMM method consistently achieves high accuracy in identifying and diagnosing faults in the switch machine. This underscores the practical applicability of the SAMM model in real-world scenarios.

5. Conclusions

This study proposes the Semi-Supervised Adaptive Matrix Machine (SAMM) model tailored to address switch machine fault diagnosis. The SAMM model features an adaptive low-rank regularizer for precise extraction of highly correlated low-rank information from matrix data and for identifying correlations between the matrix's row and column. It employs a semi-supervised learning framework that incrementally assigns pseudo-labels to unlabeled samples based on high-confidence probabilistic outputs, thereby effectively leveraging unlabeled data. An adaptive penalty term is introduced to adjust the loss penalty in response to imbalances in category sample sizes, preventing the model from being overly biased towards the majority class. Experimental validations on the switch machine current signal dataset illustrate that SAMM surpasses other baseline models in fault diagnosis accuracy. The integration of the adaptive low-rank regularizer and adaptive penalty term effectively discerns the matrix data's inherent structure. Concurrently, the semi-supervised framework augments training data through pseudo-labeling, yielding commendable classification outcomes, even with limited labeled samples.

In practical applications, the SAMM method significantly enhances railway switch machine fault diagnosis through the analysis of current signals recorded by the CSM system. This enables preventive maintenance, reduces dependency on extensive labeled datasets, lowers maintenance costs and time, and improves diagnostic accuracy by minimizing false alarms and missed detections. Additionally, the real-time monitoring capabilities of the CSM system, combined with the SAMM method, facilitate quick response to faults, thereby reducing fault handling time and ensuring the continuity and safety of railway operations.

In future research, we will focus on vibration and sound signals to explore new approaches for multimodal fault diagnosis, aiming to leverage the advantages of integrating multiple sensors. We will also investigate variations of the adaptive low-rank regularizer and extend SAMM's application to fault diagnosis and anomaly detection across diverse fields.

Author Contributions: Conceptualization, W.L., Z.X. and C.L.; methodology, W.L.; software, W.L.; validation, W.L. and M.L. ; formal analysis, W.L.; investigation, W.L.; resources, Z.X.; data curation, W.L.; writing—original draft preparation, W.L.; writing—review and editing, W.L.; visualization, W.L. and X.G.; supervision, Z.X.; project administration, M.M.; funding acquisition, Z.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China under Grant 2022YFB4300504-4 and Special Fund Project supported by Shanghai Municipal Commission of Economy and Information Technology under Grant 202201034.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: The authors would like to thank the all-round rail transit control system integrator (CASCO) for providing research data and domain knowledge support.

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

Abbreviations

The following abbreviations are used in this manuscript:

SVM	Support vector machine
CAE	Convolutional autoencoder
CNN	Convolutional neural network
SMM	Support matrix machine
MSMM	Multiclass support matrix machine

References

1. Pyrgidis, C.N. *Railway Transportation Systems: Design, Construction and Operation*; CRC Press: Boca Raton, FL, USA, 2021.
2. Loidolt, M.; Marschnig, S.; Berghold, A. Track geometry quality assessments for turnouts. *Transp. Eng.* **2023**, *12*, 100170. [CrossRef]
3. Mistry, P.; Lane, P.; Allen, P. Railway Point-Operating Machine Fault Detection Using Unlabeled Signaling Sensor Data. *Sensors* **2020**, *20*, 2692. [CrossRef] [PubMed]
4. Chen, C.; Zhu, F.; Xu, Z.; Xie, Q.; Lo, S.M.; Tsui, K.L.; Li, L. Knowledge-Informed Wheel Wear Prediction Method for High-Speed Train Using Multisource Signal Data. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 3522912. [CrossRef]
5. Luo, J.; Shao, H.; Lin, J.; Liu, B. Meta-learning with elastic prototypical network for fault transfer diagnosis of bearings under unstable speeds. *Reliab. Eng. Syst. Saf.* **2024**, *245*, 110001. [CrossRef]
6. Kim, H.; Sa, J.; Chung, Y.; Park, D.; Yoon, S. Fault diagnosis of railway point machines using dynamic time warping. *Electron. Lett.* **2016**, *52*, 818–819. [CrossRef]
7. Li, W.; Li, G. Railway's Turnout Fault Diagnosis Based on Power Curve Similarity. In Proceedings of the 2019 International Conference on Communications, Information System and Computer Engineering (CISCE), Haikou, China, 5–7 July 2019; pp. 112–115. [CrossRef]
8. Ji, W.; Cheng, C.; Xie, G.; Zhu, L.; Wang, Y.; Pan, L.; Hei, X. An intelligent fault diagnosis method based on curve segmentation and SVM for rail transit turnout. *J. Intell. Fuzzy Syst.* **2021**, *41*, 4275–4285. [CrossRef]
9. Muñoz del Río, A.; Segovia Ramirez, I.; Papaelias, M.; García Márquez, F.P. Pattern recognition based on statistical methods combined with machine learning in railway switches. *Expert Syst. Appl.* **2024**, *238*, 122214. [CrossRef]
10. Zhang, H.; Wang, Z.; Wang, N.; Long, J.; Tao, T. Fault Diagnosis of Railway Turnout Based on Random Forests. In *Proceedings of the 4th International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2019*; Qin, Y., Jia, L., Liu, B., Liu, Z., Diao, L., An, M., Eds.; Springer: Singapore, 2020; pp. 505–515.
11. Chen, C.; Li, X.; Huang, K.; Xu, Z.; Mei, M. A Convolutional Autoencoder Based Fault Detection Method for Metro Railway Turnout. *Comput. Model. Eng. Sci.* **2023**, *136*, 471–485. [CrossRef]

12. Zhang, P.; Zhang, G.; Dong, W.; Sun, X.; Ji, X. Fault Diagnosis of High-Speed Railway Turnout Based on Convolutional Neural Network. In Proceedings of the 2018 24th International Conference on Automation and Computing (ICAC), Newcastle Upon Tyne, UK, 6–7 September 2018; pp. 1–6. [CrossRef]
13. Huang, D.; Fu, Y.; Qin, N.; Gao, S. Fault diagnosis of high-speed train bogie based on LSTM neural network. *Sci. China Inf. Sci.* **2020**, *64*, 119203. [CrossRef]
14. Yan, S.; Shao, H.; Wang, J.; Zheng, X.; Liu, B. LiConvFormer: A lightweight fault diagnosis framework using separable multiscale convolution and broadcast self-attention. *Expert Syst. Appl.* **2024**, *237*, 121338. [CrossRef]
15. Shao, H.; Zhou, X.; Lin, J.; Liu, B. Few-Shot Cross-Domain Fault Diagnosis of Bearing Driven by Task-Supervised ANIL. *IEEE Internet Things J.* **2024**, *11*, 22892–22902. [CrossRef]
16. Li, X.; Li, Y.; Yan, K.; Shao, H.; Lin, J. Intelligent fault diagnosis of bevel gearboxes using semi-supervised probability support matrix machine and infrared imaging. *Reliab. Eng. Syst. Saf.* **2023**, *230*, 108921. [CrossRef]
17. Ramírez-Sanz, J.M.; Maestro-Prieto, J.A.; Arnaiz-González, Á.; Bustillo, A. Semi-supervised learning for industrial fault detection and diagnosis: A systemic review. *ISA Trans.* **2023**, *143*, 255–270. [CrossRef] [PubMed]
18. Lao, Z.; He, D.; Jin, Z.; Liu, C.; Shang, H.; He, Y. Few-shot fault diagnosis of turnout switch machine based on semi-supervised weighted prototypical network. *Knowl.-Based Syst.* **2023**, *274*, 110634. [CrossRef]
19. Shi, Z.S.; Du, Y.M.; Du, T.; Shan, G.C. The Turnout Abnormality Diagnosis with Semi-supervised Learning Method. In *Proceedings of the 4th International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2019*; Qin, Y., Jia, L., Liu, B., Liu, Z., Diao, L., An, M., Eds.; Springer: Singapore, 2020; pp. 737–746.
20. Kim, S.; Park, J.; Kim, W.; Jo, S.H.; Youn, B.D. Learning from even a weak teacher: Bridging rule-based Duval method and a deep neural network for power transformer fault diagnosis. *Int. J. Electr. Power Energy Syst.* **2022**, *136*, 107619. [CrossRef]
21. Wang, X.; Wu, Z.; Lian, L.; Yu, S.X. Debiased Learning from Naturally Imbalanced Pseudo-Labels for Zero-Shot and Semi-Supervised Learning. *arXiv* **2022**, arXiv:2201.01490.
22. Pirsiavash, H.; Ramanan, D.; Fowlkes, C. Bilinear Classifiers for Visual Recognition. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2009; Volume 22.
23. Wolf, L.; Jhuang, H.; Hazan, T. Modeling Appearances with Low-Rank SVM. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–6. [CrossRef]
24. Chen, C.; Xu, Z.; Mei, M.; Huang, K.; Lo, S.M. Fault Diagnosis Scheme for Railway Switch Machine Using Multi-Sensor Fusion Tensor Machine. *Comput. Mater. Contin.* **2024**, *79*, 4533–4549. [CrossRef]
25. Luo, L.; Xie, Y.; Zhang, Z.; Li, W.J. Support Matrix Machines. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015.
26. Zheng, Q.; Zhu, F.; Qin, J.; Heng, P.A. Multiclass support matrix machine for single trial EEG classification. *Neurocomputing* **2018**, *275*, 869–880. [CrossRef]
27. Razzak, I. Cooperative Evolution Multiclass Support Matrix Machines. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–8. [CrossRef]
28. Li, X.; Li, S.; Wei, D.; Si, L.; Yu, K.; Yan, K. Dynamics simulation-driven fault diagnosis of rolling bearings using security transfer support matrix machine. *Reliab. Eng. Syst. Saf.* **2024**, *243*, 109882. [CrossRef]
29. Liu, B.; Zhou, Y.; Liu, P.; Sun, W.; Li, S.; Fang, X. Saliency detection via double nuclear norm maximization and ensemble manifold regularization. *Know.-Based Syst.* **2019**, *183*, 12. [CrossRef]
30. Zhang, Y.; Lei, X.; Pan, Y.; Pedrycz, W. Prediction of disease-associated circRNAs via circRNA–disease pair graph and weighted nuclear norm minimization. *Knowl.-Based Syst.* **2021**, *214*, 106694. [CrossRef]
31. Jia, X.; Feng, X.; Wang, W. Adaptive regularizer learning for low rank approximation with application to image denoising. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 3096–3100. [CrossRef]
32. Zhu, W.; Peng, B. Sparse and low-rank regularized deep subspace clustering. *Knowl.-Based Syst.* **2020**, *204*, 106199. [CrossRef]
33. Xu, H.; Pan, H.; Zheng, J.; Liu, Q.; Tong, J. Dynamic penalty adaptive matrix machine for the intelligent detection of unbalanced faults in roller bearing. *Knowl.-Based Syst.* **2022**, *247*, 108779. [CrossRef]
34. Parvin, H.; Moradi, P.; Esmaeili, S.; Qader, N.N. A scalable and robust trust-based nonnegative matrix factorization recommender using the alternating direction method. *Knowl.-Based Syst.* **2019**, *166*, 92–107. [CrossRef]
35. Platt, J. Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. *Adv. Large Margin Classif.* **2000**, *10*, 61–74.
36. Wu, T.F.; Lin, C.J.; Weng, R.C. Probability Estimates for Multi-class Classification by Pairwise Coupling. *J. Mach. Learn. Res.* **2004**, *5*, 975–1005.
37. Sohn, K.; Berthelot, D.; Li, C.L.; Zhang, Z.; Carlini, N.; Cubuk, E.D.; Kurakin, A.; Zhang, H.; Raffel, C. FixMatch: Simplifying semi-supervised learning with consistency and confidence. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 596–608.

38. Cao, Y.; Sun, Y.; Xie, G.; Li, P. A Sound-Based Fault Diagnosis Method for Railway Point Machines Based on Two-Stage Feature Selection Strategy and Ensemble Classifier. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 12074–12083. [CrossRef]
39. Sun, Y.; Cao, Y.; Li, P.; Su, S. Sound Based Degradation Status Recognition for Railway Point Machines Based on Soft-Threshold Wavelet Denoising, WPD, and ReliefF. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–9. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.