TL-MC-ShuffleNetV2: A Lightweight and Transferable Framework for Elevator Guideway Fault Diagnosis

Zhiwei Zhou[1], Xianghong Deng[2], Xuwen Zheng[3], Chonlatee Photong[4]*
School of Electrical Engineering, Hunan Mechanical & Electrical Polytechnic, Changsha, 410151, Hunan, China[1, 2]
Faculty of Engineering, Mahasarakham University, Maha Sarakham, 44150, Thailand[3, 4]

***Abstract*—This study presents TL-MC-ShuffleNetV2, a lightweight and transferable fault diagnosis framework designed for elevator guideway vibration analysis. To tackle challenges such as limited labeled data and the constraints of real-time deployment, the approach integrates Variational Mode Decomposition (VMD) for multi-scale signal separation and employs a customized 1D ShuffleNetV2 backbone with multi-channel (MC) inputs. Squeeze-and-Excitation (SE) attention modules are embedded throughout the network to enhance channel-wise feature sensitivity. A transfer learning (TL) strategy is adopted, in which the model is initially trained using the Case Western Reserve University (CWRU) bearing dataset and subsequently adapted to the elevator domain by freezing early convolutional layers while fine-tuning higher-level layers. Evaluation results demonstrate that the proposed framework achieves a classification accuracy of 96.4%, alongside significantly reduced inference time and parameter complexity. Comparative and ablation experiments further validate the individual contributions of VMD preprocessing, SE modules, and transfer learning to model performance. Overall, the method exhibits strong adaptability, computational efficiency, and suitability for deployment in smart elevator monitoring systems under Industry 4.0 environments.**

Keywords*—*Transfer learning; elevator guideway; vibration signal analysis; fault diagnosis; lightweight deep neural network; squeeze-and-excitation attention; smart maintenance

I. Introduction

Elevators play a critical role in urban vertical mobility, facilitating public transportation, residential accessibility, and commercial operations [1]. With the continuous rise in urban density and high-rise development, maintaining the operational integrity and functional safety of elevator systems has become a pressing concern for building management and urban infrastructure planning [2]. However, conventional maintenance approaches—typically based on fixed schedules or reactive repairs—often struggle to detect early-stage deterioration, leading to unforeseen service interruptions, elevated maintenance expenses, and potential safety risks [3].

The emergence of Internet of Things (IoT)-enabled predictive maintenance (PdM) systems, empowered by real-time sensing and wireless communication technologies, has significantly advanced fault detection capabilities in industrial applications [4]. In particular, vibration signal analysis has proven highly effective for identifying early mechanical anomalies in components such as motors, bearings, and guide systems [5]. Alongside this development, deep learning has shown strong potential for automating the interpretation of complex, nonlinear, and nonstationary signal patterns, offering superior classification performance over traditional feature engineering-based methods [6].

Although deep learning has significantly advanced fault diagnosis techniques, its real-world application in elevator systems remains constrained by multiple unresolved challenges [7]:

Computational Constraints: Deep convolutional models require substantial computational power, making them difficult to deploy on edge devices and in latency-sensitive IoT environments [8].

Data Scarcity: Elevator-specific fault data are often scarce, particularly for early-stage or rare failures, limiting the generalization of deep models [9].

Signal Complexity: Elevator vibration signals exhibit nonstationary behavior with transient impulses and modulation effects, often buried in noise, posing challenges for reliable feature extraction [10].

To address the above challenges of nonstationary vibration signals and limited fault samples in elevator guideway diagnosis, this study builds upon our previous work on One-dimensional Convolutional Neural Network (1D-CNN) based methods using elevator vibration data [11] and proposes a lightweight and transferable fault diagnosis framework with four key processes. First, Variational Mode Decomposition (VMD) is employed to extract frequency-specific components and suppress noise [12]. Second, a tailored 1D version of ShuffleNetV2 is developed to efficiently process temporal features while significantly reducing computational complexity [13]. Third, Squeeze-and-Excitation (SE) attention modules are embedded after each network block to enhance channel-wise feature representation [14]. Fourth, a transfer learning (TL) strategy is adopted, drawing upon recent progress in TL-driven fault diagnosis [15]. Specifically, the model is initially trained using the data-rich Case Western Reserve University (CWRU) bearing dataset [16], which also with four vibration states, and then adapted to the target elevator dataset through partial parameter transfer, wherein lower convolutional layers are frozen while higher layers are updated. Compared to our earlier approach [11], the proposed method reduces the model size

Hunan Mechanical & Electrical Polytechnic Research Project (YJB202405).

while ensuring classification accuracy and model convergence, thus enhancing its applicability for deployment in resource-constrained smart elevator systems. This integrated framework is referred to as TL-MC-ShuffleNetV2, denoting the combination of Transfer Learning (TL), Multi-Channel (MC) VMD-based inputs, and a customized ShuffleNetV2 backbone.

The structure of this study is as follows: Section II surveys the relevant literature on predictive maintenance and vibration-based diagnostics. Section III details the proposed TL-MC-ShuffleNetV2 architecture for elevator fault diagnosis. Section IV describes the experimental design, comparative baselines, and ablation studies. Section V summarizes the key contributions and limitations and suggests directions for future work.

II. Related Work

A. Fault Diagnosis in Elevator Systems

As vertical transportation becomes increasingly integral to modern urban infrastructure, ensuring the operational safety and dependability of elevator systems has become a central engineering concern [17]. Fueled by Industry 4.0 and the rapid expansion of the IoT, predictive maintenance (PdM) has gained prominence as an advanced alternative to conventional reactive or periodic maintenance strategies [18]. In this context, vibration-based fault diagnosis plays a pivotal role in identifying anomalies such as misalignment, wear, or guideway deformation before catastrophic failures occur [19].

Extensive research has investigated the use of data-driven methods for elevator fault detection. For instance, Lv et al. developed a GAN-based model to synthesize vibration data for intelligent fault classification in elevators [20]. Qiu et al. introduced a hybrid approach combining the improved Aquila optimizer (IAO) and XGBoost to enhance fault pattern recognition from vibration signals [21]. Liu et al. proposed a framework integrating ensemble empirical mode decomposition (EEMD) with support vector machines (SVMs) to identify faults in traction systems [22], while Song et al. incorporated EMD with neural network architectures to detect anomalies in elevator control systems [23].

Despite these developments, practical deployment faces three core challenges: 1) real-world elevator vibration signals are nonstationary and influenced by diverse operating conditions, making them difficult to model using conventional linear analysis methods [24]. 2) The scarcity of labeled fault data—especially for rare or incipient cases—substantially limits the effectiveness of supervised learning approaches [25]. 3) Standard deep learning models are computationally demanding and ill-suited for real-time application in edge-based IoT monitoring systems with limited hardware resources [26].

These limitations motivate the need for lightweight, adaptive models that can operate effectively with limited labeled data and computational capacity, while still providing reliable performance on noisy, complex vibration signals.

B. Deep Learning and Lightweight Models for Vibration-Based Fault Diagnosis

Deep learning, especially various CNNs, has demonstrated strong effectiveness in performing automated fault recognition by learning hierarchical features directly from unprocessed sensor data. Unlike traditional time- or frequency-domain methods such as RMS analysis or Fourier transforms, CNNs can derive distinctive patterns without requiring prior domain knowledge, thereby enhancing scalability and adaptability across diverse fault scenarios [27].

However, conventional CNN architectures—despite their effectiveness—are often computationally intensive and unsuitable for real-time fault monitoring in embedded IoT environments. To address this issue, lightweight CNNs such as MobileNet, EfficientNet, and ShuffleNet have been developed, leveraging techniques like depthwise separable convolutions, grouped operations, and channel shuffling to minimize parameter overhead and reduce model complexity and inference latency [28]. ShuffleNetV2, in particular, offers a compelling trade-off between speed and accuracy, making it a promising candidate for industrial fault diagnosis applications where low-resource deployment is essential [29].

Beyond efficiency, recent studies emphasize the importance of attention mechanisms in improving model interpretability and performance. Representative modules, such as Squeeze-and-Excitation Networks (SE-Net) and the Convolutional Block Attention Module (CBAM), enable the model to selectively concentrate on diagnostically significant frequency components while attenuating irrelevant noise and redundant information. For instance, CBAM has been demonstrated to enhance classification performance in machinery fault datasets by dynamically reweighting spatial and channel-specific features [30].

Overall, combining lightweight architectures with attention mechanisms enhances the feature representation of nonstationary signals while ensuring computational feasibility, especially in IoT-enabled predictive maintenance systems.

C. Transfer Learning for Small-Sample Industrial Scenarios

In many real-world scenarios—particularly in elevator guideway diagnostics—acquiring large, labeled datasets is often impractical due to cost limitations, safety concerns, or the infrequent nature of fault events. Transfer learning (TL) provides a powerful solution by allowing a model trained on a source domain with ample data to be adapted to a target domain with limited samples [31]. This approach has demonstrated strong applicability across diverse industrial settings.

For example, Yang et al. employed a sparse autoencoder-based TL method to enhance bearing fault classification on limited datasets [32], while Azad et al. proposed dynamic modeling strategies that simulate data to support domain adaptation using convolutional networks [33]. In wind turbine maintenance, parameter-based TL combined with autoencoders has enabled models trained on one machine to generalize effectively to others with similar operating characteristics [34].

These studies indicate that effective domain transfer enables successful cross-domain adaptation. For elevator systems, limited sample learning and fine-tuning of pre-trained models offer a promising approach for addressing data scarcity challenges.

III. Methodology

A. Variational Mode Decomposition for Multiscale Fault Feature Separation

Elevator guideway vibration signals are inherently nonstationary, exhibiting complex modulation patterns due to fluctuating operational loads and external disturbances. To effectively capture fault-relevant features across diverse frequency bands, this study employs Variational Mode Decomposition (VMD). VMD adaptively partitions the signal into a set of band-limited intrinsic mode functions (IMFs), each aligned with a distinct spectral component, thereby improving the interpretability and diagnostic precision of nonstationary signal analysis [35].

The decomposition process is formulated as a constrained variational optimization problem, aiming to reduce the aggregated bandwidth of all modes while ensuring that their summation reconstructs the original signal [36]. Mathematically, the VMD objective is defined as:

$$min_{u_k,\omega_k}\left\{\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\} \quad (1)$$

subject to

$$\sum_k u_k(t)=f(t) \quad (2)$$

where, $u_k(t)$ denotes the k-th IMF, w_k represents its center frequency, ∂_t is the first-order time derivative, $\delta(t)$ is the Dirac delta function, and $f(t)$ denotes the original signal.

In this study, the first five IMFs (as shown in Fig. 1)—capturing the most representative low- to mid-frequency fault features—are selected to construct a five-channel input for the subsequent feature extraction network. By leveraging the superior mode separation and noise resilience of VMD, this strategy enhances the model's sensitivity to subtle fault-related variations, thereby improving its applicability to real-world elevator vibration analysis under uncertain operating conditions.

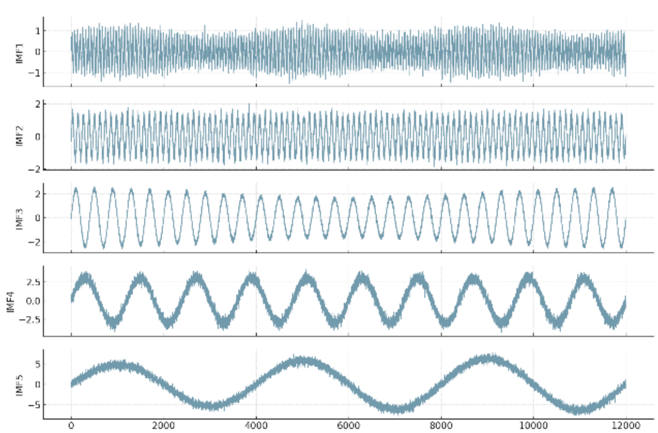

Fig. 1. Variational Mode Decomposition (VMD) processing results.

B. Modified 1D ShuffleNetV2 Architecture for Feature Extraction

To effectively extract discriminative features from the multi-channel IMF signals generated by VMD, a customized one-dimensional version of ShuffleNetV2 is designed. Originally developed for efficient image classification, the ShuffleNetV2 architecture is adapted here to handle sequential vibration signals by converting all convolutional operations into their 1D forms [37]. This enables the model to process temporal input while retaining the core advantages of ShuffleNetV2—namely, lightweight design and efficient feature utilization.

The modified 1D ShuffleNetV2 (as shown in Fig. 2) consists of a series of Shuffle Units, each of which employs a dual-branch design [38]. For units with stride = 1, the input is first split into two equal channel groups. One branch directly propagates its features without modification, while the other branch undergoes a sequence of three operations:

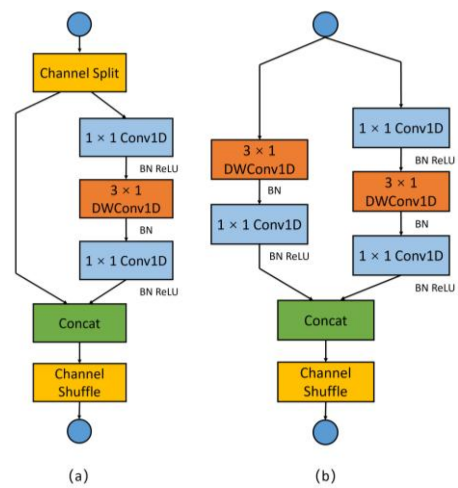

Fig. 2. Modified 1D ShuffleNetV2 architecture: a) Stride = 1 block; b) Stride = 2 block.

1) A 1×1 pointwise convolution followed by batch normalization and ReLU activation,

2) A 3×1 depthwise convolution with subsequent batch normalization, and

3) A concluding 1×1 pointwise convolution, also accompanied by batch normalization and ReLU activation.

After feature extraction, both branches are concatenated and passed through a channel shuffle operation, which promotes inter-group feature interaction and enhances representation diversity.

For units with stride 2, both branches perform independent downsampling. Each branch executes a sequence of depthwise separable and pointwise convolutions, followed by concatenation and channel shuffling. This design strategy reduces spatial resolution while increasing output channel capacity. The channel split step is intentionally excluded in this configuration to prevent structural fragmentation and enhance feature utilization.

This modular design enables the network to strike a balance between computational cost and expressive power, thereby ensuring its applicability to fault diagnosis tasks in IoT environments with limited resources. Furthermore, by leveraging grouped convolutions and the channel shuffle mechanism, the model effectively extracts fine-grained temporal features from multiple IMFs while preserving a lightweight parameter footprint.

C. Equations Channel Attention Enhancement with SE Module

To enhance the network's focus on informative features and suppress irrelevant activations, a lightweight channel attention mechanism—Squeeze-and-Excitation (SE) module—was incorporated into the proposed architecture. This mechanism facilitates channel-wise recalibration by explicitly capturing inter-channel dependencies, thereby improving the network's representational capacity in a computationally efficient manner.

The SE module functions through two primary stages [39]: squeeze and excitation. In the squeeze stage, global contextual information is aggregated from each channel using global average pooling. This operation produces a channel-wise descriptor that reflects the global distribution of activations. During the excitation phase, two fully connected layers with a bottleneck structure and nonlinear activations (ReLU and Sigmoid) are employed to model channel-wise dependencies and generate attention weights. These weights are then applied to reweight the original feature maps through channel-wise multiplication, enabling the model to highlight informative channels while suppressing less relevant ones.

Mathematically, for an intermediate feature map $X \in R^{C \times T}$, where C and T represent the number of channels and temporal length, respectively, the SE module first derives a channel-wise descriptor via global average pooling [40]:

$$z_c = \frac{1}{T}\sum_{t=1}^{T} X_{c,t}, \quad c = 1, 2, \ldots, C \tag{3}$$

The attention weights are then generated through a gating mechanism:

$$\mathbf{s} = \sigma\left(\mathbf{W}_2 \cdot \delta\left(\mathbf{W}_1 \cdot \mathbf{z}\right)\right) \tag{4}$$

where, $\delta(\cdot)$ and $\sigma(\cdot)$ represent the ReLU and Sigmoid functions, respectively, and $W1$, $W2$ are learnable weight matrices. The final recalibrated output is obtained by channel-wise multiplication:

$$\tilde{X}_c = s_c \cdot X_c \tag{5}$$

As illustrated in Fig. 3, the SE module is integrated into each Shuffle Block of the Modified 1D ShuffleNetV2 architecture. Specifically, it is inserted after the depthwise convolution (DWConv1D) and before the final pointwise convolution (1×1 Conv1D). This placement enables the attention mechanism to refine intermediate features that already encode localized spatial and inter-channel information. The recalibrated features are subsequently propagated through the remaining layers of the Shuffle Block, followed by concatenation and channel shuffling operations.

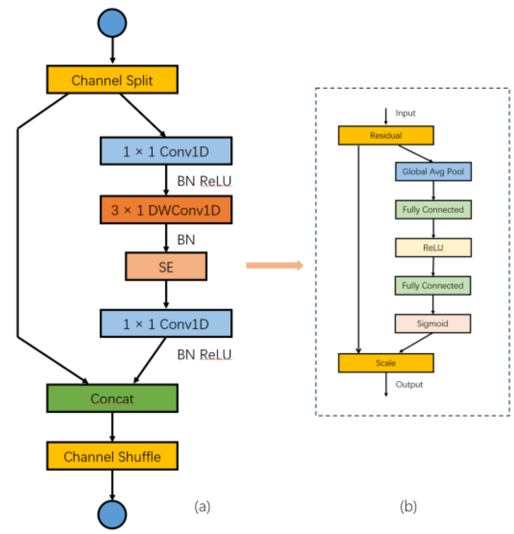

Fig. 3. Integration of SE module into modified 1D ShuffleNetV2 architecture (a), and structure and operation of SE block (b).

The SE module was inserted in both Stage 1 (stride = 2) and Stage 2 (stride = 1) Shuffle Blocks, ensuring that channel attention is applied during both the downsampling and feature extraction processes. This integration strategy enhances the network's ability to adaptively weight different channels according to task-relevant information, leading to improved fault diagnosis performance in the subsequent classification stage.

D. Transfer Learning Strategy for Fault Diagnosis

To address the limited availability of labeled elevator guideway fault data, a transfer learning (TL) strategy is employed. The training procedure consists of two key stages:

1) Pre-training stage. The proposed framework, incorporating VMD-enhanced multi-channel IMFs, is input into the modified 1D ShuffleNetV2 augmented with SE modules. Pre-training is conducted on the CWRU bearing dataset that also has four states, enabling the model to capture generalized vibration features from the rich source domain data to compensate for the insufficient data in the target domain.

2) Fine-tuning stage. The pre-trained model is then adapted to the elevator guideway dataset. Lower layers are frozen to retain transferable low-level representations, while upper layers are selectively fine-tuned based on the target domain data.

This two-phase learning paradigm enables efficient domain adaptation with minimal supervision, ensuring that transferable vibration characteristics are preserved while tailoring feature representations to the specific requirements of the target dataset.

E. Overview of the Proposed TL-MC-ShuffleNetV2 Framework

The proposed TL-MC-ShuffleNetV2 framework (illustrated in Fig. 4) integrates all preceding components into a cohesive fault diagnosis pipeline tailored for elevator guideway vibration analysis. The process begins by applying VMD to the raw vibration signals, generating multiple IMFs that preserve localized frequency information. These IMFs are organized as MC inputs and fed into a customized one-dimensional ShuffleNetV2 backbone optimized for temporal feature extraction with low computational cost. To enhance channel-wise feature discrimination, SE modules are selectively embedded after each ShuffleNetV2 block. To address the scarcity of labeled elevator data, a transfer learning strategy is employed. The model is initially pre-trained on the CWRU bearing dataset to capture transferable mechanical fault representations. It is then fine-tuned using the target elevator dataset by freezing the initial convolutional layers and updating the higher ones. After domain adaptation, the model is evaluated on previously unseen elevator vibration data, with minor parameter refinements applied to further enhance performance. This modular and efficient design promotes robust cross-domain generalization while maintaining a lightweight architecture, making it well-suited for real-world deployment in smart elevator systems.

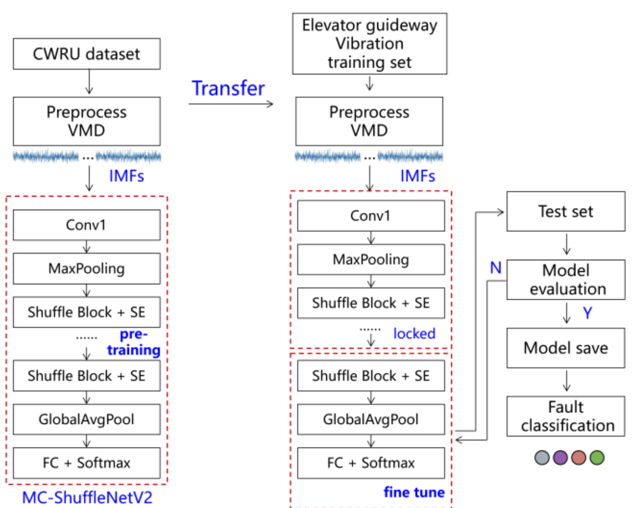

Fig. 4. Structural diagram of the proposed TL-MC-ShuffleNetV2-based fault diagnosis framework.

IV. Experiments and Result Analysis

A. Description of Data Sets

In this study, two vibration datasets were utilized to support the transfer learning framework—one for source domain pre-training and the other for target domain fine-tuning. The source domain employs a publicly available bearing dataset, while the target domain comprises elevator guideway vibration signals collected under real-world operational conditions.

1) Source domain CWRU bearing dataset. As a widely recognized benchmark for mechanical fault diagnosis, CWRU bearing dataset [16] contains vibration signals acquired from an experimental motor testbed equipped with SKF 6205-2RS bearings, simulating various fault types.

For pre-training, four fault categories were selected: normal, inner race, outer race, and ball fault. The raw time-series signals, sampled at 12 kHz, were segmented using a 1024-point sliding window with 50% overlap, yielding 1400 samples per class (5600 total). An 80/20 split was adopted for training and testing purposes.

2) Target domain elevator guideway dataset. The target dataset was collected via an industrial elevator IoT platform, with tri-axial accelerometers mounted at the top-center of elevator cars to capture guideway vibration signals. The sampling frequency was 50 Hz, and four operating states were considered: normal, bending, misalignment, and step fault.

Using the same segmentation approach as the source domain, 300 samples were extracted for each class, resulting in a total of 1200 samples. These samples were split into 80/20 training and testing sets for target domain fine-tuning and evaluation. Details of the dataset composition and sample allocation are summarized in Table I. The table adapts the structure and sample configuration from our prior work [11].

TABLE I. DATASETS USED FOR TRANSFER LEARNING

	Dataset	Category	Label	Train sample	Test sample
Source domain (pre-train)	CWRU bearing	Normal	0	1120	280
		Ball fault	1	1120	280
		Inner fault	2	1120	280
		Outer fault	3	1120	280
Target domain	Elevator guideway	Normal	0	240	60
		Bending	1	240	60
		Misalignment	2	240	60
		Step	3	240	60

B. Experimental Settings and Model Training

Model training was carried out on a workstation equipped with an Intel i7-12700K processor, 32 GB of memory, and an RTX 3080 GPU. Python 3.9 and PyTorch 1.13 were used throughout the implementation.

The proposed MC-ShuffleNetV2 architecture accepts five IMFs as MC inputs, obtained through VMD. The network integrates a modified 1D ShuffleNetV2 backbone for lightweight feature extraction, with SE modules inserted after each Shuffle Block to enhance channel attention.

Table II summarizes the architecture and parameter settings of the MC-ShuffleNetV2 with SE model prior to transfer learning (TL). The TL pipeline described earlier in the methodology is subsequently applied. with pre-training conducted on the CWRU bearing dataset and fine-tuning performed using the elevator guideway dataset. The training strategy parameters are summarized in Table III.

TABLE II. ARCHITECTURE DETAILS OF MC-SHUFFLENETV2 WITH SE

Layer	Output Size	Kernel Size	Stride	Repeat	Description
Input	5×1024×1	–	–	–	5 IMFs
Conv1	512×32	3×1	2	1	1D convolution
MaxPooling	256×32	3×1	2	1	Downsampling
Shuffle Block 1 + SE	128×64	3×1	2	1	Stage 1
Shuffle Block 2 + SE	128×64	3×1	1	4	Stage 2
GlobalAvgPool	1×64	–	–	–	Global Average pooling
FC + Softmax	1×4	–	–	–	Output for 4 categories

TABLE III. TRAINING STRATEGY PARAMETERS

Parameter Name	Optimizer	Initial Learning Rate	Learning Rate Decay	Loss Function	Batch Size	Epochs	Cross-Validation Strategy
Value	Adam	0.0005	0.5 / 10 epochs	CrossEntropy	32	100	5-fold cross-validation

C. Evaluation Metrics

To comprehensively assess the diagnostic capability of the proposed TL-MC-ShuffleNetV2 framework, several widely used classification metrics were employed to evaluate both overall and class-specific performance.

Accuracy reflects the share of total predictions that match the actual class labels:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

Precision evaluates the proportion of correct positive predictions:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{7}$$

Recall indicates the extent to which actual faults are successfully flagged:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{8}$$

F1-Score offers a balanced measure for classification evaluation:

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

where, TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. The Confusion Matrix offers an intuitive visualization of classification outcomes across categories, facilitating the identification of misclassification trends.

D. Results and Analysis

1) Convergence curves of training. Fig. 5 illustrates the training accuracy and loss curves recorded over 100 epochs during cross-validation. The network demonstrates rapid convergence within the first 20 epochs, with training accuracy exceeding 95% and training loss stabilizing thereafter. Despite early convergence, the model was trained for 100 epochs to maintain consistency in the evaluation process and support fair comparison across experiments.

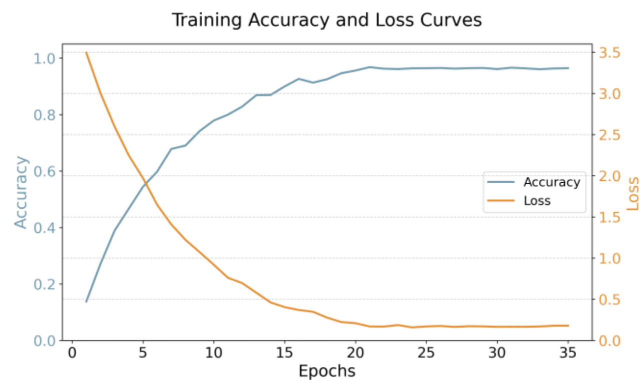

Fig. 5. Training accuracy and loss curves.

2) Confusion matrix analysis. The confusion matrix for the testing set is presented in Fig. 6. The proposed model achieved 100% accuracy in identifying normal samples. For fault categories, minor misclassifications were observed between bending and misalignment types, likely due to similar dynamic behaviors in their vibration signals. However, step faults exhibited distinct characteristics and were more easily separated. These results indicate that the model is capable of distinguishing subtle differences among fault types with high reliability.

3) Feature visualization. To further assess the feature discriminative capacity of the proposed TL-MC-ShuffleNetV2 model, the t-distributed stochastic neighbor embedding (t-SNE) technique was employed to map high-dimensional features—extracted prior to the classification layer—onto a two-dimensional space for visualization. The resulting projection is illustrated in Fig. 7.

As illustrated, samples from class 0 (normal) are distinctly clustered with clear boundaries, indicating the model's ability to robustly identify normal operational conditions. Class 3 (step fault) also forms a relatively compact and isolated cluster, suggesting reliable recognition. In contrast, class 1 (bending) and class 2 (misalignment) show a degree of overlap in the projected space, reflecting a partial confusion between these two types of faults—consistent with the observations from the confusion matrix.

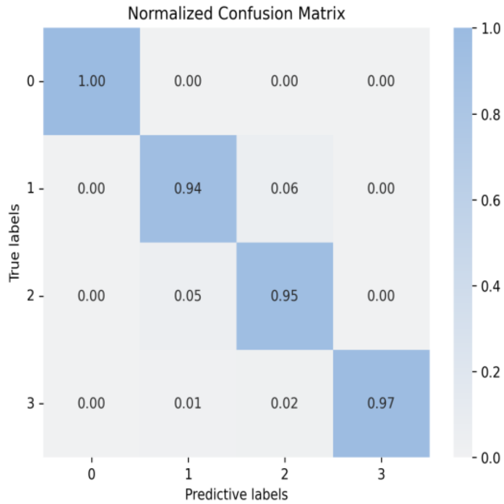

Fig. 6. Normalized confusion matrix (test set).

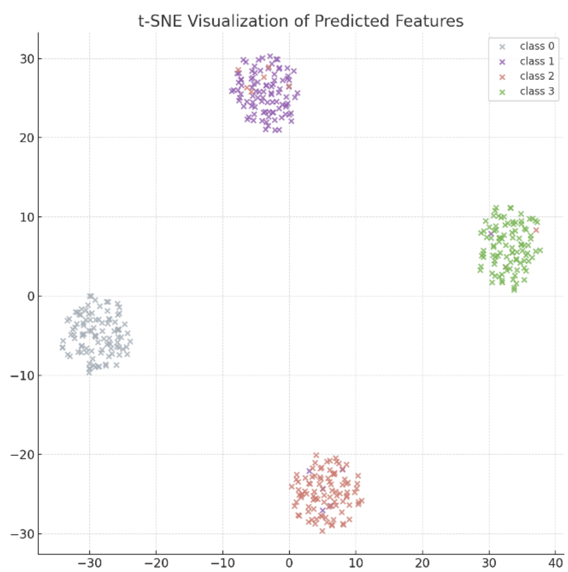

Fig. 7. t−SNE feature visualization.

E. Comparison with Other Models

To thoroughly assess the effectiveness of the proposed TL-MC-ShuffleNetV2 framework, comparative experiments were conducted against several representative baseline models, including 1D-CNN, ResNet18-1D, MobileNetV2-1D, and EfficientNet-B0-1D. The evaluation was carried out using five standard classification metrics—Accuracy, Precision, Recall, F1-score—and the Number of Parameters (Params). These indicators jointly reflect both the predictive performance and computational efficiency of each model. Table IV presents a detailed comparison of the results on the elevator guideway fault dataset.

The results show that TL-MC-ShuffleNetV2 achieves the best balance between accuracy and model compactness, outperforming deeper models such as ResNet18-1D and EfficientNet-B0-1D, while maintaining significantly fewer parameters.

TABLE IV. PERFORMANCE COMPARISON OF DIFFERENT MODELS

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Params (M)
1D-CNN	91.2	90.5	90.7	90.6	0.8
ResNet18-1D	93.5	92.8	93.0	92.9	11.2
MobileNetV2-1D	94.3	93.6	93.9	93.7	2.5
EfficientNet-B0-1D	95.1	94.5	94.8	94.6	4.0
TL-MC-ShuffleNetV2	**96.4**	**95.8**	**96.1**	**95.9**	**1.7**

To further illustrate the trade-off between accuracy and computational efficiency, Fig. 8 presents a comparative bar chart highlighting the Accuracy (%) and Inference Time (ms) of each model.

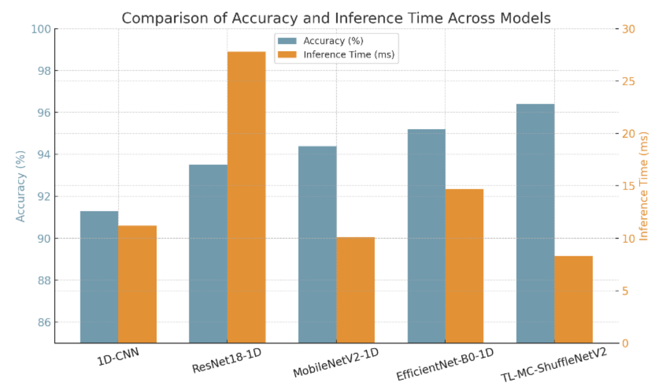

Fig. 8. Comparison of accuracy and inference time across models.

The comparison results reveal that although ResNet18-1D achieves satisfactory accuracy, its inference time is the longest due to its high parameter complexity. In contrast, the TL-MC-ShuffleNetV2 model delivers competitive classification performance while significantly reducing inference latency to 7.8 ms, highlighting its potential for real-time fault diagnosis in resource-constrained IoT environments.

These findings further validate that the proposed TL-MC-ShuffleNetV2 not only excels in feature learning but also offers exceptional efficiency, making it a strong candidate for practical elevator guideway fault detection applications.

F. Ablation Study

To assess the individual contribution of each major component in the proposed TL-MC-ShuffleNetV2 framework, three controlled ablation variants are constructed:

w/o VMD: Removes VMD, raw vibration signals are directly input into the network.

w/o SE: Removes all SE modules to evaluate the effect of channel attention.

w/o TL: Disables TL, the model is trained solely on the target domain without pretraining.

All configurations are trained under consistent settings with 5-fold cross-validation. The classification accuracy over epochs is illustrated in Fig. 9.

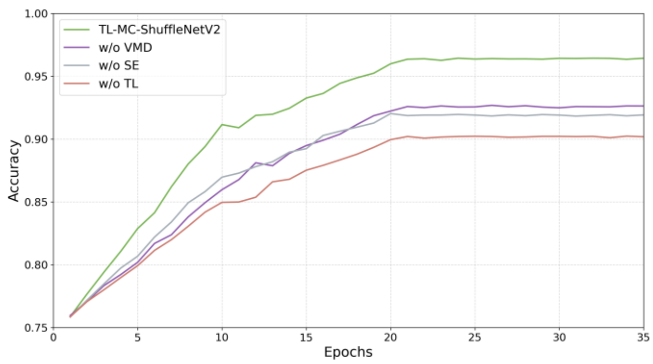

Fig. 9. Ablation study results (accuracy over training epochs).

The comparative results indicate the following:

- The w/o TL configuration shows the lowest final accuracy and slowest convergence, suggesting that TL significantly enhances generalization in limited data settings.
- Excluding VMD also results in reduced performance, confirming the effectiveness of multi-scale signal decomposition in extracting fault-relevant features.
- The removal of SE modules leads to a moderate decline, indicating that channel-wise recalibration contributes to better feature representation.

These observations validate the importance of all three components—VMD, SE, and TL—in achieving the superior performance of the proposed model.

V. CONCLUSION

This study presents TL-MC-ShuffleNetV2, an efficient and lightweight fault diagnosis framework tailored for elevator guideway vibration analysis. Addressing the challenges of small-sample learning and edge deployment, the proposed method integrates VMD for multiscale signal decomposition, a modified multi-channel 1D ShuffleNetV2 for feature extraction, an SE attention module to enhance channel-wise representation, and a transfer learning strategy that learns generic features of vibration faults from similar domain datasets to compensate for data scarcity in elevator guideway fault diagnosis.

Extensive experiments validate the framework's superiority in both accuracy and computational efficiency. Achieving 96.4% classification accuracy, the model significantly reduces inference time and parameter overhead, outperforming representative baselines such as 1D-CNN, ResNet18-1D, and MobileNetV2-1D. Ablation studies further confirm the individual contributions of VMD, SE modules, and TL strategy to overall performance. The framework demonstrates high accuracy and lightweight parameters, making it well-suited for resource-constrained IoT deployments in elevator fault monitoring.

Despite demonstrating strong diagnostic accuracy and adaptability under typical low-speed operating conditions for a specific elevator brand, the proposed method has not yet been thoroughly validated across a broader range of elevator models or varying operational speeds. Future studies will aim to enhance the framework's generalizability under more diverse and dynamic conditions. Additionally, the framework will be expanded to address fault diagnosis in other key elevator components, contributing to a more holistic condition monitoring solution. Further directions may also consider integrating few-shot learning techniques to alleviate current reliance on both labeled data and transfer learning strategies, improving adaptability in real-world deployment scenarios.

REFERENCES

[1] A. C. H. J. Thebuwena, S. M. S. M. K. Samarakoon, and R. M. C. Ratnayake, "Optimization of energy consumption in vertical mobility systems of high-rise office buildings: A case study from a developing economy," Energy Efficiency, vol. 17, no. 6, p. 68, 2024/08/14 2024, doi: 10.1007/s12053-024-10246-5.

[2] Y. Li, W. Zheng, and Q. Zhou, "Knowledge-driven urban innovation: dynamics of elevator installation in aging residential communities," Journal of the Knowledge Economy, pp. 1-45, 2024.

[3] J. West, M. Siddhpura, A. Evangelista, and A. Haddad, "Improving Equipment Maintenance—Switching from Corrective to Preventative Maintenance Strategies," Buildings, vol. 14, no. 11, p. 3581, 2024.

[4] D. Dhinakaran, S. Edwin Raja, R. Velselvi, and N. Purushotham, "Intelligent IoT-Driven Advanced Predictive Maintenance System for Industrial Applications," SN Computer Science, vol. 6, no. 2, p. 151, 2025.

[5] I. Bagri, K. Tahiry, A. Hraiba, A. Touil, and A. Mousrij, "Vibration Signal Analysis for Intelligent Rotating Machinery Diagnosis and Prognosis: A Comprehensive Systematic Literature Review," Vibration, vol. 7, no. 4, pp. 1013-1062, 2024.

[6] G. S. Kollu and J. Kittur, "Integrating Deep Learning Into Mechanical Engineering: A Systematic Review of Applications and Educational Implications," Computer Applications in Engineering Education, vol. 33, no. 3, p. e70048, 2025.

[7] A. Saeed, M. A. Khan, U. Akram, W. J. Obidallah, S. Jawed, and A. Ahmad, "Deep learning based approaches for intelligent industrial machinery health management and fault diagnosis in resource-constrained environments," Scientific Reports, vol. 15, no. 1, p. 1114, 2025.

[8] H.-I. Liu et al., "Lightweight deep learning for resource-constrained environments: A survey," ACM Computing Surveys, vol. 56, no. 10, pp. 1-42, 2024.

[9] Y. F. Zhao, J. Xie, and L. Sun, "On the data quality and imbalance in machine learning-based design and manufacturing—A systematic review," Engineering, 2024.

[10] A. Achilleos, Q. Leclere, and J. Antoni, "Feature Extraction in nonstationary conditions," in Surveillance, Vibrations, Shock and Noise, 2023.

[11] Z. Zhou, X. Zheng, T. Chen, and C. Photong, "A transfer learning method based on MC-1DCNN for elevator guideway fault diagnosis," Edelweiss Applied Science and Technology, vol. 9, no. 3, pp. 633-650, 2025.

[12] G. Wang, L. Sun, A. Wang, J. Jiao, and J. Xie, "Lithium battery remaining useful life prediction using VMD fusion with attention mechanism and TCN," Journal of Energy Storage, vol. 93, p. 112330, 2024.

[13] H. Yang, J. Liu, G. Mei, D. Yang, X. Deng, and C. Duan, "Research on real-time detection method of rail corrugation based on improved ShuffleNet V2," Engineering Applications of Artificial Intelligence, vol. 126, p. 106825, 2023.

[14] C. Zeng, Y. Zhao, Z. Wang, K. Li, X. Wan, and M. Liu, "Squeeze-and-excitation self-attention mechanism enhanced digital audio source recognition based on transfer learning," Circuits, Systems, and Signal Processing, pp. 1-33, 2024.

[15] I. Misbah, C. K. Lee, and K. L. Keung, "Fault diagnosis in rotating machines based on transfer learning: Literature review," Knowledge-Based Systems, vol. 283, p. 111158, 2024.

[16] D. Neupane and J. Seok, "Bearing Fault Detection and Diagnosis Using Case Western Reserve University Dataset With Deep Learning Approaches: A Review," IEEE Access, vol. 8, pp. 93155-93178, 2020, doi: 10.1109/ACCESS.2020.2990528.

[17] M.-L. Siikonen, "Current and future trends in vertical transportation," European Journal of Operational Research, 2024.

[18] A. A. Murtaza, A. Saher, M. H. Zafar, S. K. R. Moosavi, M. F. Aftab, and F. Sanfilippo, "Paradigm shift for predictive maintenance and condition monitoring from Industry 4.0 to Industry 5.0: A systematic review, challenges and case study," Results in Engineering, p. 102935, 2024.

[19] R. Kumar et al., "A state-of-the-art review on the misalignment, failure modes and its detection methods for bearings," Mapan, vol. 38, no. 1, pp. 265-274, 2023.

[20] X. Lv, Z. Lu, Z. Huang, and Z. Wei, "Improving fault diagnosis in elevator systems with GAN-based synthetic data," International Journal of Sensor Networks, vol. 47, no. 1, pp. 26-35, 2025.

[21] C. Qiu, L. Zhang, M. Li, P. Zhang, and X. Zheng, "Elevator Fault Diagnosis Method Based on IAO-XGBoost under Unbalanced Samples," Applied Sciences, vol. 13, no. 19, 2023, doi: https://doi.org/10.3390/app131910968.

[22] H. Liu, L. Yang, and Y. Wei, "Elevator traction machine fault diagnosis based on improved CEEMD and SVM Algorithm," in 2022 Global Reliability and Prognostics and Health Management (PHM-Yantai), 2022: IEEE, pp. 1-6, doi: https://doi.org/10.1109/PHM-Yantai55411.2022.9941777.

[23] Q. Song, Q. Song, L. Xiao, H. Jiang, and L. Li, "State Diagnosis of Elevator Control Transformer over Vibration Signal Based on MEA - BP Neural Network," Shock and Vibration, vol. 2021, no. 1, p. 9755094, 2021, doi: https://doi.org/10.1155/2021/9755094.

[24] E. I. Vlachou, V. I. Vlachou, D. E. Efstathiou, and T. S. Karakatsanis, "Overview of IoT Security Challenges and Sensors Specifications in PMSM for Elevator Applications," Machines, vol. 12, no. 12, p. 839, 2024.

[25] J. Yu and Y. Zhang, "Challenges and opportunities of deep learning-based process fault detection and diagnosis: a review," Neural Computing and Applications, vol. 35, no. 1, pp. 211-252, 2023.

[26] F. Oliveira, D. G. Costa, F. Assis, and I. Silva, "Internet of Intelligent Things: A convergence of embedded systems, edge computing and machine learning," Internet of Things, p. 101153, 2024.

[27] T. Mahesh, S. Chandrasekaran, V. A. Ram, V. V. Kumar, V. Vivek, and S. Guluwadi, "Data-driven intelligent condition adaptation of feature extraction for bearing fault detection using deep responsible active learning," IEEE Access, vol. 12, pp. 45381-45397, 2024.

[28] S. M. Raza, S. M. H. Abidi, M. Masuduzzaman, and S. Y. Shin, "Survey on the Application with Lightweight Deep Learning Models for Edge Devices," Authorea Preprints, 2025.

[29] R. Khanam, M. Hussain, R. Hill, and P. Allen, "A comprehensive review of convolutional neural networks for defect detection in industrial applications," IEEE Access, 2024.

[30] B. Li, B. Tang, L. Deng, and J. Wei, "Joint attention feature transfer network for gearbox fault diagnosis with imbalanced data," Mechanical Systems and Signal Processing, vol. 176, p. 109146, 2022.

[31] P. Singhal, R. Walambe, S. Ramanna, and K. Kotecha, "Domain adaptation: challenges, methods, datasets, and applications," IEEE access, vol. 11, pp. 6973-7020, 2023.

[32] J. Yang, G. Xie, and Y. Yang, "Fault diagnosis of rotating machinery using denoising-integrated sparse autoencoder based health state classification," Ieee Access, vol. 11, pp. 15174-15183, 2023.

[33] S. Azad and M. T. Ameli, "A domain adaptation - based convolutional neural network incorporating data augmentation for power system dynamic security assessment," The Journal of Engineering, vol. 2024, no. 7, p. e12400, 2024.

[34] Y. Li, W. Jiang, G. Zhang, and L. Shu, "Wind turbine fault diagnosis based on transfer learning and convolutional autoencoder with small-scale data," Renewable Energy, vol. 171, pp. 103-115, 2021, doi: https://doi.org/10.1016/j.renene.2021.01.143.

[35] R. Pavithra and P. Ramachandran, "Non-stationary signal analysis utilizing VMD and EMD algorithm for multiple fault classification for industrial applications," Journal of Intelligent & Fuzzy Systems, pp. JIFS-237546, 2024.

[36] Y. He, W. Wang, M. Li, and Q. Wang, "A short-term wind power prediction approach based on an improved dung beetle optimizer algorithm, variational modal decomposition, and deep learning," Computers and Electrical Engineering, vol. 116, p. 109182, 2024.

[37] Z. Luo, H. Tan, X. Dong, G. Zhu, and J. Li, "A fault diagnosis method for rotating machinery with variable speed based on multi-feature fusion and improved ShuffleNet V2," Measurement Science and Technology, vol. 34, no. 3, p. 035110, 2022.

[38] W. Jiang, Z. Qi, A. Jiang, S. Chang, and X. Xia, "Lightweight Network Bearing Intelligent Fault Diagnosis Based on VMD-FK-ShuffleNetV2," Machines, vol. 12, no. 9, p. 608, 2024.

[39] X. Jin, Y. Xie, X.-S. Wei, B.-R. Zhao, Z.-M. Chen, and X. Tan, "Delving deep into spatial pooling for squeeze-and-excitation networks," Pattern Recognition, vol. 121, p. 108159, 2022.

[40] M. Z. Nisar, M. S. Ibrahim, M. Usman, and J.-A. Lee, "A lightweight deep learning model for automatic modulation classification using residual learning and squeeze–excitation blocks," Applied Sciences, vol. 13, no. 8, p. 5145, 2023.