applied sciences

MDPI

Article

A Text Generation Method Based on a Multimodal Knowledge Graph for Fault Diagnosis of Consumer Electronics

Yuezhong Wu [1], Yuxuan Sun [1], Lingjiao Chen [1,*], Xuanang Zhang [1] and Qiang Liu [2]

[1] College of Railway Transportation, Hunan University of Technology, Zhuzhou 412007, China; wuyuezhong@hut.edu.cn (Y.W.); m22081101018@stu.hut.edu.cn (Y.S.); m23081100015@stu.hut.edu.cn (X.Z.)
[2] School of Computer, Hunan University of Technology, Zhuzhou 412007, China; liuqiang@hut.edu.cn
* Correspondence: chenlingjiao@hut.edu.cn

Abstract: As consumer electronics evolve towards greater intelligence, their automation and complexity also increase, making it difficult for users to diagnose faults when they occur. To address the problem where users, relying solely on their own knowledge, struggle to diagnose faults in consumer electronics promptly and accurately, we propose a multimodal knowledge graph-based text generation method. Our method begins by using deep learning models like the Residual Network (ResNet) and Bidirectional Encoder Representations from Transformers (BERT) to extract features from user-provided fault information, which can include images, text, audio, and even olfactory data. These multimodal features are then combined to form a comprehensive representation. The fused features are fed into a graph convolutional network (GCN) for fault inference, identifying potential fault nodes in the electronics. These fault nodes are subsequently fed into a pre-constructed knowledge graph to determine the final diagnosis. Finally, this information is processed through the Bias-term Fine-tuning (BitFit) enhanced Chinese Pre-trained Transformer (CPT) model, which generates the final fault diagnosis text for the user. The experimental results show that our proposed method achieves a 4.4% improvement over baseline methods, reaching a fault diagnosis accuracy of 98.4%. Our approach effectively leverages multimodal fault information, addressing the challenges users face in diagnosing faults through the integration of graph convolutional network and knowledge graph technologies.

Keywords: multimodal; graph convolutional network; knowledge graph; text generation; consumer electronics

check for updates

Citation: Wu, Y.; Sun, Y.; Chen, L.; Zhang, X.; Liu, Q. A Text Generation Method Based on a Multimodal Knowledge Graph for Fault Diagnosis of Consumer Electronics. *Appl. Sci.* **2024**, *14*, 10068. https://doi.org/10.3390/app142110068

Academic Editor: Mohamed Benbouzid

Received: 25 September 2024
Revised: 29 October 2024
Accepted: 31 October 2024
Published: 4 November 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

With the advancement of information and communication technology, the era of the Internet of Things (IoT) has arrived. As a typical representative of this era, consumer electronics continue to develop rapidly [1,2]. Certain consumer electronics, including smart TVs, smart air conditioners, and smart light bulbs, demonstrate substantial potential in the consumer market [3]. As consumer electronics increasingly advance toward greater intelligence, their automation and complexity continue to rise, which also gradually increases the likelihood of malfunctions. When faults occur, they can lead to significant economic losses and, in severe cases, even result in injury or loss of life. Table 1 lists the occurrence frequency of common faults in consumer electronics. Faults in consumer electronics are more challenging to detect than in most mechanical systems and structures due to their micro- to nano-scale complex architectures. Faults within electronic products do not necessarily lead to outright failure or loss of specified electrical performance or functionality, making it difficult to quantify the progression from fault occurrence to complete failure [4]. As shown in Figure 1, latent faults in consumer electronics are often difficult for users to identify.

Table 1. The occurrence frequency of the most common faults in electronic devices.

Fault Type	Description	Occurrence Frequency (%)
Battery Issues	Shortened battery life or inability to charge	20–30%
Screen Problems	Display malfunctions or touch sensitivity loss	15–20%
Overheating	Device overheating, leading to performance drop	10–15%
Software Crashes	Frequent crashes of system or applications	10–12%
Power Supply Issues	Malfunctions in charger or power management	8–10%
Audio Problems	Speaker or headphone jack malfunctions	5–8%
Mechanical Failures	Buttons or switches becoming unresponsive	5–7%

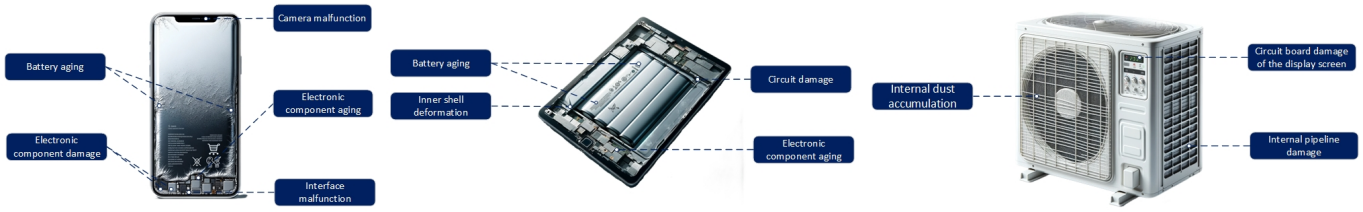


Figure 1. Diagram of latent faults in consumer electronics.

Fault diagnosis methods for consumer electronics primarily include user experience-based diagnostics and model-based diagnostics [5]. While these methods are capable of handling fault diagnosis tasks in most areas, the accuracy and stability of the diagnosis are often susceptible to external influences. Experience-based diagnostics are highly subjective, as users' experience, knowledge, and observational abilities are limited, which can lead to inaccuracies in the diagnosis and even the potential to overlook critical fault information [6]. Traditional model-based fault diagnosis methods often utilize fault information insufficiently and incompletely. Typically, models rely on single-modal information for product fault diagnosis or use knowledge graphs alone to query product faults. These approaches often struggle to produce accurate diagnostic results when faced with complex fault scenarios in consumer electronics. Methods that use multimodal information for fault diagnosis are significantly superior to traditional approaches. For example, Kullu et al. [7] proposed a deep learning-based multimodal sensor fusion method that detects equipment faults by integrating data from various sensors. The study specifically highlights the advantages of multimodal fusion over single-sensor data in detecting equipment anomalies, as multimodal fusion greatly enhances diagnostic accuracy. Nan et al. [8] proposed a multimodal learning framework within an IoT environment that improves the accuracy of equipment fault prediction by integrating both low-quality and high-quality monitoring data. Yang et al. [9] proposed an attention-based fault diagnosis method for multimodal data in chemical processes. By applying attention mechanisms to multimodal data, the model enhances its ability to select and integrate features from different modalities, thereby improving fault diagnosis accuracy. Compared to single-modal methods, multimodal fault diagnosis approaches offer significant advantages.

In summary, diagnosing faults in consumer electronics can be achieved by integrating multiple modalities of information with knowledge graphs. We propose a multimodal knowledge graph-based method for fault diagnosis that effectively combines different types of information to achieve reliable results. Compared to traditional methods, our approach fully utilizes multimodal data, enabling comprehensive and multi-level diagnosis, which addresses the challenge of accurately identifying complex faults in consumer electronics. The main contributions of this paper are as follows:

(1) We propose an innovative fault diagnosis method for consumer electronics that fully utilizes multimodal fault information. By integrating multiple types of data during a fault, multimodal features are extracted and used for diagnosis. This approach significantly enhances the efficiency of fault diagnosis, making it easier for users to manage their devices.

(2) We design a fault text generation model based on a multimodal knowledge graph specifically for consumer electronics. By integrating a semi-supervised multimodal graph convolutional network, the model improves the ability to capture correlations between fault data, ensuring the accuracy of the generated fault diagnosis text.
(3) Experimental results show that our model outperforms baseline models in F1 score and accuracy.

The rest of the paper is organized as follows: Section 2 reviews related work, Section 3 describes the model's architecture and methods, Section 4 details the experiments and results analysis, Section 5 provides a summary, and Section 6 outlines future work.

2. Related Work

Given the importance of reliability in consumer electronics, research on fault diagnosis for these products has become a key area of development. Pecht et al. [10] proposed a health management method based on a physics-of-failure model, integrating sensor data with the model to assess the health status of electronic products. Kumar et al. [11] introduced a unique hybrid prognostics and health management method, combining data-driven and physics-of-failure models for fault diagnosis in electronics. Kumar et al. [12] proposed a fault diagnosis method for electronic products based on the Mahalanobis distance. Stetter et al. [13] were the first to attempt formulating general, effective guidelines for fault diagnosis in mechanical and electrical products to achieve efficient and effective diagnostics. Cai et al. [14] addressed the issue of performance degradation in electronic products over time by proposing a fault diagnosis method for electronic systems with transient and intermittent faults, using a dynamic Bayesian network. Nguyen et al. [15] proposed using historical data from unobserved components in electronic devices to reduce the search space for potential faulty components, thereby diagnosing faults in electronic devices more accurately. Arcos et al. [16] investigated how users diagnose faults in household appliances and qualitatively analyzed fault diagnosis in consumer electronics such as kitchen blenders, vacuum cleaners, and refrigerators. Their analysis showed that consumers identify faults by focusing on five main conditions. Based on these findings, they proposed design recommendations for household electronics to facilitate user-driven fault diagnosis. Lin [17] proposed a sensor fault diagnosis network for consumer electronics based on the transformer architecture, applying artificial intelligence technology to the field of fault diagnosis for consumer electronics. Jin et al. [18] introduced trace ratio linear discriminant analysis to handle high-dimensional non-Gaussian fault data for dimensionality reduction and fault classification, validating the proposed fault diagnosis method with motor-bearing data containing single-point faults and generalized roughness faults. Souri et al. [19] presented a hybrid machine learning-based fault prediction model using a multilayer perceptron and particle swarm optimization algorithm, applying the model to fault diagnosis in electronic products. Liu et al. [20] proposed an online sensor fault detection framework. Fernandes et al. [21] summarized fault diagnosis methods for household appliances. Song et al. [22] reviewed common fault diagnosis methods for power electronic circuits in consumer electronics, discussing intelligent fault diagnosis technology for power electronic circuits from four perspectives: building a comprehensive data information database, detecting faults in multiple transmission lines, accurately simulating circuit operation, and diagnosing main circuit faults. Lall et al. [23] conducted anomaly detection on electronic products subjected to shock and vibration. Marijan et al. [24] proposed a fully automated system for detecting functional faults in televisions, with the proposed method outperforming other algorithms. Li et al. [25] proposed an intelligent diagnostic solution for digital medical devices. Liang et al. [26] addressed the ambiguity and uncertainty of fault probability in electronic products by introducing fuzzy set theory into fault tree analysis for electronic products, scientifically analyzing various fuzzy information encountered in electronic product faults and providing examples of applying fuzzy fault tree analysis in electronic product diagnostic systems. However, these fault diagnosis methods for consumer electronics still require further research and refinement.

Multimodal knowledge graphs are widely applied in many fields. Wu et al. [27] proposed an industrial equipment fault diagnosis model based on a multimodal knowledge graph, addressing several challenges faced in industrial equipment fault diagnosis. Zhang et al. [28] proposed a knowledge graph technology framework utilizing knowledge graphs and natural language processing to support the intelligent management of transformers. This approach addresses the difficulties of using unstructured text data and the shallow application of equipment knowledge during the inspection process of transformer equipment in power supply companies. Deng et al. [29] constructed an event logic knowledge graph for fault diagnosis of robotic transmission systems, providing decision support for autonomous fault diagnosis in these systems. Multimodal knowledge graphs are widely applied in the industrial sector, but their application in fault diagnosis of consumer electronics is relatively limited. The use of multimodal graph neural networks for diagnosing faults in consumer electronics is still in the exploratory stage.

Text generation technology has recently garnered significant attention from academia and industry, with applications in fields such as media publishing, biomedical sciences, and advertising copy. Luo et al. [30] applied natural language processing techniques to the biomedical field. Lebret et al. [31] proposed a neural model for concept-to-text generation and applied this approach to the biographical domain, where it significantly outperformed other language models. Ji et al. [32] provided an extensive overview of the hallucination problem in natural language generation. Xu et al. [33] utilized large-scale language models and incorporated external knowledge bases to add control over text generation. While intelligent text generation technology has demonstrated great potential, it still faces challenges such as poor controllability, lack of interpretability, and high resource consumption.

3. Method Implementation

When a fault occurs in a consumer electronic device, users can upload diagnostic information, including fault images, descriptions (either as voice or text), and odor data from sensors. If only a single type of information is provided, such as an image, the model helps users add more information, for example, by providing a text input box or a voice module. This ensures comprehensive data collection for the model. The collected information is then analyzed using graph neural networks and a multimodal knowledge graph to identify the root cause of the fault. Finally, the results are transformed into a user-friendly report using a text generation model, making the diagnosis easy to understand.

Following this approach, after collecting user-provided fault information, the model extracts features from various modalities of fault data. For images, the model first preprocesses the fault image by resizing and normalizing it. The image is then input into a ResNet network, where it passes through convolutional and pooling layers to obtain an initial feature map. This feature map is further refined through residual blocks to extract deeper features. Finally, after global pooling and a fully connected layer, the image feature vector is obtained.

For text data, the model preprocesses the fault description using the BertTokenizer to tokenize the text, adding special markers and positional encodings to create an input vector. This input vector is then fed into BERT's encoding layer, where self-attention mechanisms generate context-aware feature vectors. Feature pooling is applied to produce the text feature vector.

For audio data, the model preprocesses the fault audio by adjusting the sampling rate and segmenting the audio into frames. These frames are then input into a Long Short-Term Memory (LSTM) model, which processes each frame sequentially to generate a probability distribution over text. This distribution is passed through a decoder to produce output text, which is then input into the BERT model to generate the audio feature vector.

For other modalities, data dimensionality expansion is applied before extracting feature vectors. For example, for odor data collected by sensors, the model maps the one-dimensional odor data to a higher-dimensional feature space, capturing the characteristics of the olfactory data more effectively and resulting in an olfactory feature vector.

After extracting feature vectors from image, text, audio, and olfactory modalities, the model uses a modality fusion module to integrate these feature vectors. The fusion module employs a contrastive learning strategy that maximizes the similarity between related modality features while minimizing the similarity between unrelated ones, producing a fused multimodal feature vector. This process ensures the seamless integration of information from different modalities, resulting in a comprehensive input for subsequent fault detection and diagnosis.

The multimodal feature vector is then fed into GraphSAGE for inference. At this stage, the model maps the multimodal feature vector as nodes within a graph and establishes adjacency relationships to construct the graph. GraphSAGE is then used to sample neighboring nodes and aggregate their features. Through multi-level feature aggregation, each layer of GraphSAGE updates the node feature vectors, allowing each node representation to include both its own features and those of its neighbors. This process enables the node features to capture information from nodes at greater distances. By sampling and aggregating with GraphSAGE, the model utilizes the multimodal feature vectors to infer the fault diagnosis results.

The fault results obtained from GraphSAGE are input into a pre-constructed multimodal knowledge graph, where reasoning is used to identify fault points and corresponding diagnostic solutions for consumer electronics. The model first uses cosine similarity to match the node feature vectors inferred by GraphSAGE with entity vectors in the knowledge graph, determining their semantic relevance to find the most similar entity. Next, reasoning is performed based on the matched entity and its associated relationships within the knowledge graph. During this process, semantic matching templates are also used to enhance inference accuracy, ultimately identifying the cause of the fault and providing corresponding recommendations for resolution. The primary knowledge base within the multimodal knowledge graph consists of entities such as devices, fault types, and diagnostic solutions, capturing relationships between faults and devices, as well as between devices and diagnostic pathways. By leveraging knowledge graph reasoning, the model efficiently identifies faults and offers users actionable diagnostic solutions.

Based on the results of multimodal knowledge graph reasoning, a fine-tuned CPT text generation model is used to generate a fault diagnosis report, providing users with specific fault descriptions and handling recommendations. The model takes the reasoning results as input for the fine-tuned CPT model. Based on the input fault type and related information, the CPT model generates diagnostic text, which includes two main components: a fault description detailing the current fault status of the device, and handling recommendations to guide the user on effectively resolving the issue. The overall structure of the model is illustrated in Figure 2.

3.1. Feature Extraction and Graph Neural Network

Multimodal feature extraction and fusion are crucial steps in deep learning for processing multimodal data. In this study, the model extracts feature vectors from various modalities of fault information provided by the user, including image, text, audio, and odor. These feature vectors capture the rich semantic information of each modality. The model then fuses these multimodal features to create a comprehensive representation.

To achieve effective fusion of multimodal features (such as image, text, audio, and odor), this study draws inspiration from the work of Contrastive Language-Image Pre-Training (CLIP) [34] and designs a contrastive learning framework. By maximizing the similarity between correctly matched modality pairs and minimizing the similarity between mismatched pairs, this framework enables robust multimodal feature fusion. Specifically, as shown in Equation (1), cosine similarity is first used to compute the similarity between

different modalities, where f_x and f_y represent the feature vectors of any two modalities, such as the image–text or text–audio modality pairs. Subsequently, a contrastive loss L, as defined in Equation (2), is used to optimize the fusion of multimodal features, where f_{other} denotes the feature vector of any other modality. With this loss function, the model can maximize the similarity of correctly paired modalities and minimize the similarity of incorrectly paired modalities during training. Throughout this process, highly similar modality feature vectors gradually draw closer together, while less similar or unrelated modality feature vectors are pushed apart. Finally, the contrastive losses L for each modality pair are summed to obtain the total loss. By minimizing this total loss, the model achieves effective fusion of multimodal features, resulting in a unified multimodal feature vector.

$$sim(f_x, f_y) = \frac{f_x \cdot f_y}{\| f_x \| \| f_y \|} \tag{1}$$

$$L(x, y) = -\log \frac{\exp(sim(f_x, f_y))}{\exp(sim(f_x, f_{\text{other}}))} \tag{2}$$

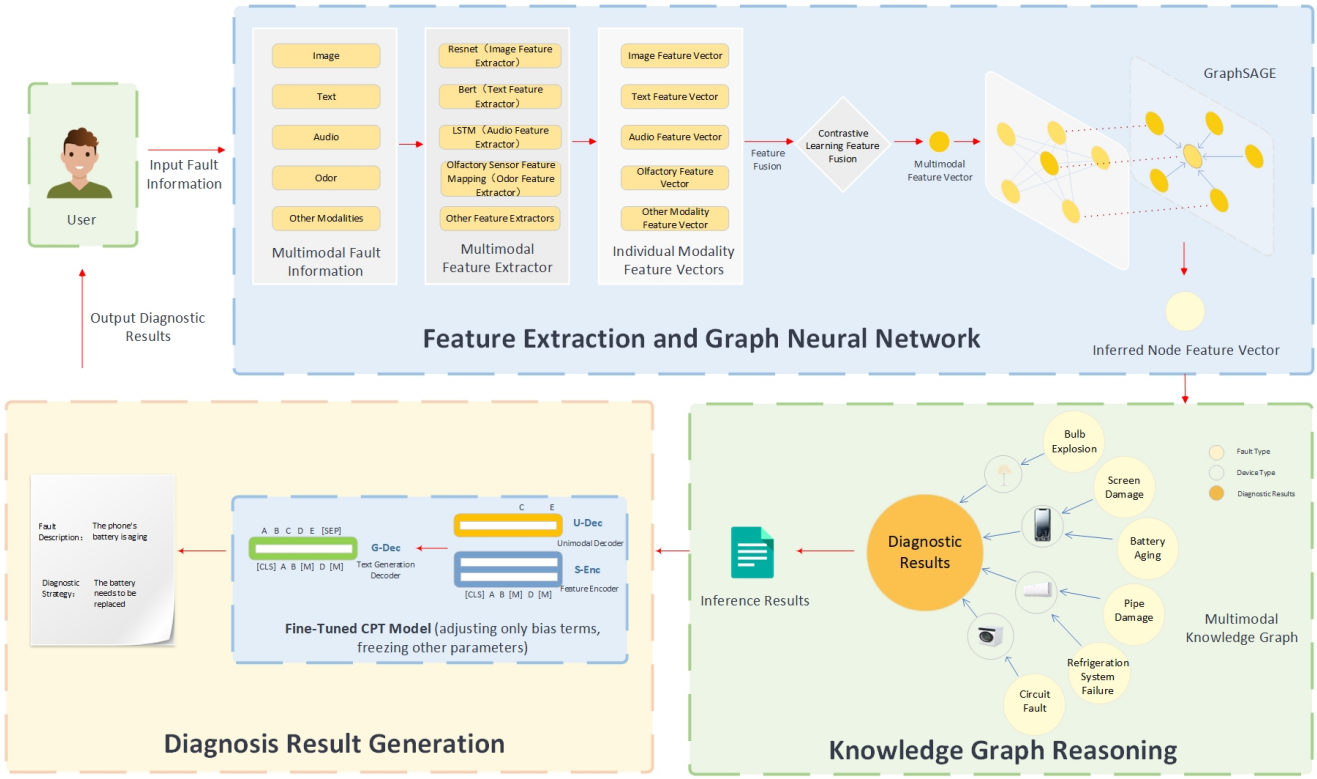


Figure 2. Overall framework of the model.

After extracting the multimodal feature vectors, this study uses GraphSAGE for fault inference. GraphSAGE is a spatial-based GCN model that performs operations like direct aggregation on node information. Compared to traditional GCN models, GraphSAGE offers a more efficient feature aggregation approach.

In GCN, feature aggregation for each node takes a global approach, requiring information from all neighboring nodes, which increases computational and memory costs as the graph grows. In contrast, GraphSAGE uses a 'sample-aggregate' method, randomly sampling only a subset of each node's neighbors. This significantly reduces memory and computational burdens, making GraphSAGE more suitable for large-scale graphs.

GraphSAGE also surpasses GCN in its aggregation capabilities. While GCN typically relies on basic aggregation methods like mean or weighted aggregation, GraphSAGE provides multiple options, including mean aggregation, LSTM-based aggregation, and others.

These options capture richer relationships among neighboring nodes, allowing for more flexible feature representation tailored to different graph structures and data characteristics.

The network framework of GraphSAGE is illustrated in Figure 3a, where 'K' represents the number of network layers. The central red node is the target node, while the surrounding nodes are neighbors at different hop levels: yellow nodes are 1-hop neighbors, green nodes are 2-hop neighbors, and blue nodes are 3-hop neighbors. By incorporating information from these multi-hop neighbors, the model progressively updates the feature representation of the central node.

The GraphSAGE training process consists of two main steps: sampling and aggregation. GraphSAGE uses a neighbor sampling approach, optimizing the global sampling method of GCN by determining the central node and then randomly sampling its neighbors. Using this method, GraphSAGE samples and aggregates features of neighboring nodes layer by layer to update the central node's representation.

As shown in Figure 3b, GraphSAGE's neighbor sampling method is illustrated. At each layer, neighboring nodes are randomly sampled, with sampled nodes (in orange) used for feature aggregation, while unsampled nodes remain white. This sampling mechanism reduces computational complexity while maintaining efficient aggregation, making GraphSAGE suitable for consumer electronics fault diagnosis tasks.

Figure 3c depicts the aggregation and inference steps of GraphSAGE. The model aggregates features from various layers and inputs them into the inference network, where three different aggregators (Aggregator1, Aggregator2, Aggregator3) fuse the information to predict the label of the central node. In the figure, 'MF' represents the multimodal feature vector, and 'Label' denotes the node's feature. By leveraging GraphSAGE's sampling and aggregation, the model uses multimodal feature vectors to infer fault diagnosis results.

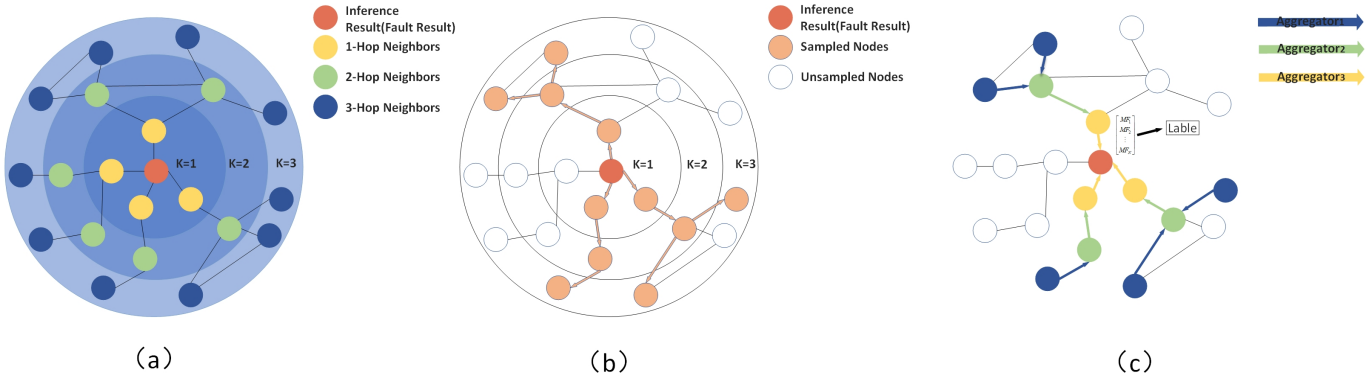


Figure 3. (**a**) GraphSAGE network architecture. (**b**) Sampled neighborhood for GraphSAGE. (**c**) Feature aggregation and label prediction.

$$J_{\mathcal{G}} = -log(\sigma(z_u^{\top} z_v)) - Q \cdot \mathbb{E}_{v_n \sim p_n}(v) log(\sigma(-z_u^{\top} z_{v_n})) \tag{3}$$

Equation (3) represents the loss function of GraphSAGE, where z_u and z_v denote the embedding vectors of two nodes u and v. These embeddings are node representations computed by GraphSAGE through aggregating neighboring node features. z_{v_n} represents the embedding vector of a negative sample v_n, which is a randomly selected other node used for contrastive learning to help the model differentiate between positive and negative samples. Q denotes the negative sampling multiplier, indicating how many negative samples should be paired with each positive sample. $\mathbb{E}_{v_n \sim p_n}(v)$ represents the expectation of the sampling probability distribution p_n for negative sample nodes v_n. By defining this loss function, GraphSAGE can efficiently accomplish the inference task.

3.2. Knowledge Graph Reasoning

The consumer electronics fault knowledge graph is constructed using a top-down approach, given the domain-specific nature of the data. High-confidence information, such

as images, user descriptions, and audio related to faults in consumer electronics, is collected from the internet and incorporated into the knowledge base to describe fault types and their relationships. The data are stored in a graph database in a networked format to build the knowledge graph, enabling associative querying similar to human thought processes. This knowledge graph plays a crucial role in enhancing the model's intelligence and enabling autonomous fault diagnosis.

The fault information obtained from GraphSAGE is input into the knowledge graph, where the knowledge graph extracts a feature vector to obtain the entity information. This entity vector serves as a text feature vector, which is then mapped to entities and relationships within the knowledge graph. Cosine similarity is used to effectively align these relationships during the mapping process.

Using cosine similarity, the entity text vector is matched and mapped to entities in the knowledge graph. Leveraging fault entities, relationships, and tail entities that represent fault handling suggestions, a triplet is formed. This triplet is then processed through a semantic matching template to generate fault diagnosis results and handling recommendations, as shown in Figure 4.

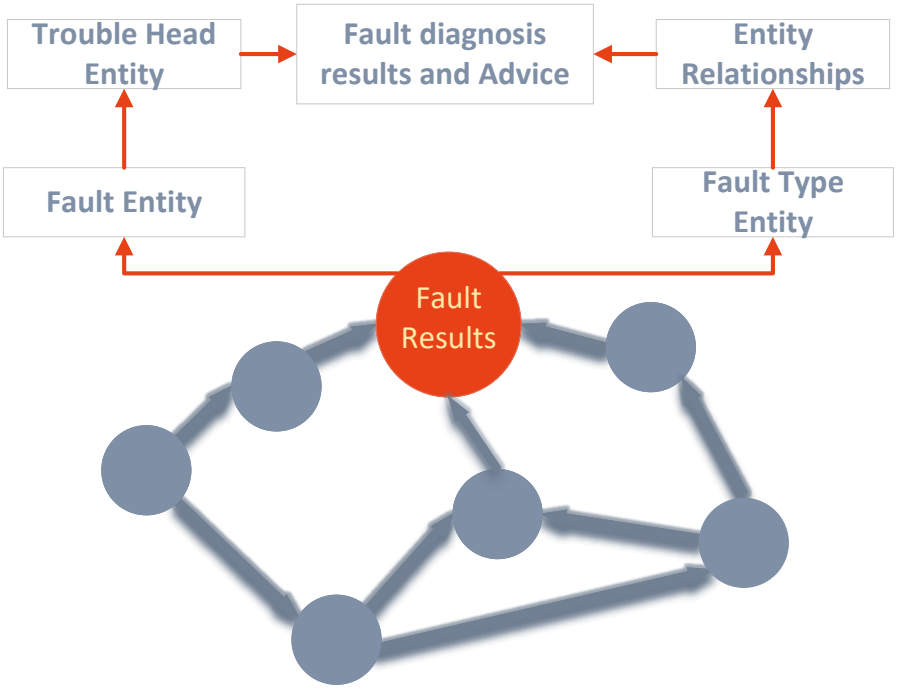


Figure 4. Multimodal knowledge graph reasoning.

3.3. Diagnosis Result Generation

After obtaining the fault diagnosis results from the knowledge graph, the fault text generation model uses a Bidirectional and Auto-Regressive Transformers (BART) language model to generate diagnostic text. BART employs a neural machine translation approach based on the Transformer architecture, adding a noise function to perturb the input text and generate corrupted sequences. Through sequence-to-sequence learning, the model is trained to reconstruct the original text. This enables BART to handle noisy text effectively, demonstrating high efficiency in text generation tasks.

For our work, we use the CPT model [35] for text generation tasks. Under a multi-task learning framework, the CPT model is pre-trained on two tasks, enhancing its performance in Chinese text comprehension and generation. To further optimize the model for the specific fault text generation task in consumer electronics, we fine-tuned the CPT model using the BitFit approach. By adjusting only a subset of the model's bias terms, BitFit enables efficient fine-tuning with reduced parameter updates, significantly enhancing both

inference efficiency and text generation capability. This fine-tuned model is better suited for fault diagnosis tasks.

3.4. Algorithm Design

Summarize the overall algorithm based on the approach presented in this paper. After preparing the dataset, use feature extraction networks to extract feature vectors of different modalities. These feature vectors are then fused into a multimodal feature vector using a contrastive learning strategy. GraphSAGE is subsequently used for inference, and the inference result must exceed a set threshold to be considered reliable data. In this paper, the threshold setting is discussed. By reviewing relevant literature, we initially chose a lower threshold (e.g., 50%) for experimentation and found that the model generated too many false positives during fault diagnosis, leading to poor performance in accuracy, recall, and F1 score. When we selected a higher threshold (e.g., 70%), it missed some real faults in consumer electronics, resulting in false negatives and affecting diagnostic efficiency. We then adjusted the threshold and found that setting it at 60% achieved optimal results in terms of accuracy, recall, and F1 score. Based on experimental observations and a review of relevant literature, we consider inference results with a probability greater than 60% as reliable data, while results with a probability below 60% are regarded as unreliable data. The fault results obtained by GraphSAGE are compared with entities in the multimodal knowledge graph based on similarity, yielding the final fault description entity. Through reasoning based on relationships between entities, the fault diagnosis result is derived. This diagnosis result is then input into the fine-tuned CPT model to generate the final fault diagnosis text, which is subsequently presented to the user. The algorithm design of the model is shown in Algorithm 1.

Algorithm 1 Modal Fusion and Reasoning Process

Data: Multiple modal data
Result: Knowledge graph inference results

initialization;
while ResNet and Bert obtain multiple modal information $M_m \in M$ **do**
    multimodal vector $= Softmax(sim(f_x, f_y))$
end while
GraphSAGE get multimodal vector;
$\mathcal{J}_\mathcal{G}$ get multimodal vector;
get failure results;
if failure probability $\geq$ 60% **then**
    knowledge graph get failure result;
else
    no failure
end if
if failure results text feature $=$ head entity **then**
    unrecognizable;
else
    semantic matching templates;
end if
get knowledge graph inference results R;
CPT get R;
fault description and solution;

4. Experiment

4.1. Experimental Configuration

The experimental setup of this study is shown in Table 2.

Table 2. Experimental configuration table.

Configure	System Environment
RTX4060(12 G) I7	Pytorch1.13 Python3.7 CUDA11.6 CUDNN8.4.0 Neo4j4.2.2 windows11

4.2. Data Collection

The dataset used in this study includes image, text, audio, and olfactory data. The fault image dataset was collected through web scraping, starting with 1024 images that were manually filtered down to 850, covering faults in devices like smartphones, tablets, and smart air conditioners. The fault text dataset was constructed using 'How to Diagnose and Fix All Electronic Devices' [36]. Entity recognition, entity linking, and relationship extraction were performed on electronic fault descriptions from the book. Identified entities and relationships were vectorized and stored in a knowledge base to create the text dataset.

Fault audio data were collected from the internet, while the olfactory dataset was developed based on national standards and environmental simulations. It includes five representative fault-related odors commonly found in consumer electronics: CF_2Cl_2, Hg, H_2S, CO, and O_3. These odors were simulated, and their concentration values processed to align dimensionally with image features. The features were then sequentially fused, resulting in 4550 sets of characteristics. An example of the consumer electronics fault dataset is shown in Figure 5.

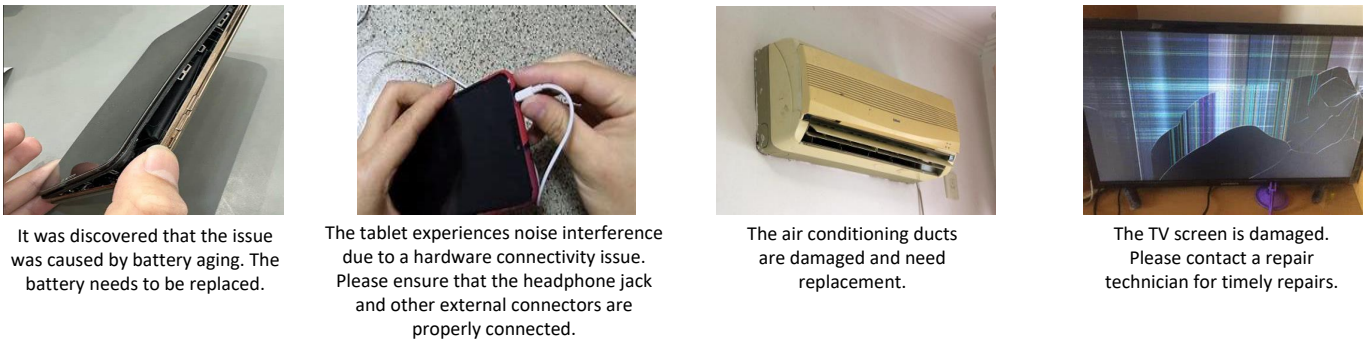


Figure 5. Displaying dataset samples.

4.3. Experimental Indicators

The evaluation metrics for this experiment include precision (P), recall (R), F1 score ($F1$), mean average precision (mAP), $Rouge-N$, $Rouge-L$, and $BLEU$.

P represents the proportion of samples classified as positive that are actually positive. R indicates the proportion of truly positive samples that are correctly classified as positive. The $F1$ score takes both P and R into account and is their harmonic mean. The mAP value is the mean of average precision and is commonly used to evaluate the model's performance in fault detection. $Rouge-N$ is a metric for evaluating the quality of text generation by the model. As shown in Equation (4), $\text{Count}_{match}(gram_n)$ represents the number of matching n-gram between the candidate text and the reference text.

$$Rouge-N = \frac{\sum_{S\in\{Reference\}} \sum_{gram_n\in S} Count_{match}(gram_n)}{\sum_{S\in\{Reference\}} \sum_{gram_n\in S} Count(gram_n)} \quad (4)$$

$Rouge-L$ is a metric used to evaluate the similarity between generated text and reference text. It is based on the longest common subsequence (LCS) and calculates similarity by finding the longest subsequence shared between the generated and reference texts. $Rouge-L$ can measure the sequential consistency and retention of key information between the texts. The calculation formula for $Rouge-L$ is shown in Equation (5), where $\mathrm{LCS}(X,Y)$ represents the length of the longest common subsequence between the generated text X and the reference text Y. $|X|$ and $|Y|$ are the lengths of the generated and reference texts, respectively.

$$Rouge-L=\frac{LCS(X,Y)}{\max(|X|,|Y|)} \tag{5}$$

BLEU is a metric used to evaluate the difference between the generated text and the reference text. It is chosen to assess the quality of text generation due to its effectiveness in measuring the degree of match between the generated text and the reference text. As shown in Equations (6) and (7), $max_{j\in m}hk(s_{ij})$ represents the maximum number of occurrences of an n-gram across multiple reference sentences, while $\sum_i\sum_k min(h_k(c_i),max_{j\in m}hk(s_{ij}))$ represents the minimum number of occurrences of an n-gram between the generated sentence and the reference sentence. BP stands for the brevity penalty factor.

$$P_n=\frac{\sum_i\sum_k min(h_k(c_i),max_{j\in m}hk(s_{ij}))}{\sum_i\sum_k min(h_k(c_i))} \tag{6}$$

$$BLEU=BP\times exp(\sum_{n=1}^{N}W_n log(P_n)) \tag{7}$$

4.4. Comparison Models

We selected several state-of-the-art models for performance comparison, including GraphSAGE, GNN, GCN, and GAT. GNN uses deep learning to directly process graph-structured data, providing a meaningful baseline. GCN operates on graphs, effectively utilizing structural information. GAT incorporates an attention mechanism, enabling efficient neighbor node aggregation, making it a suitable baseline for comparison.

In the consumer electronics fault text generation phase, we compare our model with the mT5, RoBERTa, and BART language models. The mT5 model is effective in multilingual environments and performs well across various benchmarks. The RoBERTa model excels in tasks like sentiment analysis and text generation. The BART model demonstrates strong pre-training and multitask learning capabilities for text generation. Comparing our model with these advanced models is highly significant.

4.5. Results Analysis

To validate the performance of the proposed multimodal GraphSAGE model for consumer electronics fault diagnosis, we compared it with other advanced models on the same dataset. The experimental results are presented in Table 3, where '-SI' represents the results using single-modality image features, and '-M' represents the results using multimodal features.

Table 3. Performance comparison of models.

Algorithm	mAP (%)	F1 (%)	P (%)
GNN-SI	44.2	45.2	51.1
GCN-SI	51.7	48.1	54.5
GAT-SI	57.5	49.9	57.9
GraphSAGE- SI	56.2	51.8	60.1
GraphSAGE- M	59.8	50.2	62.3

As shown in Table 3, when using single-modal information alone, we observe that GAT achieves a higher mAP score than GraphSAGE. However, after incorporating multimodal information, GraphSAGE surpasses GAT in terms of mAP. This might be due to the fact that under single-modal conditions, the structural complexity of GraphSAGE, which is an advantage in handling complex graph relationships, might not capture information as comprehensively as GAT when processing single-modal data. In contrast, when multiple modalities are integrated, GraphSAGE demonstrates superior performance over other comparison models. This improvement could be attributed to its adaptability and scalability, allowing it to leverage shared graph structures effectively and capture the relational information across different modalities as node information increases.

In terms of F1 score, the GraphSAGE model shows better performance under single-modal conditions, which may be due to the model's recall performance influencing the result. The precision trends align with the mAP results, showing improved performance with multimodal information. This consistency across metrics highlights the model's capability to leverage multimodal data for more accurate fault diagnosis in consumer electronics.

The combination of knowledge graphs with graph neural networks enables the fault detection model for consumer electronics to achieve improved performance and enhanced fault description capabilities. Compared to traditional fault detection methods or single-modal deep learning-based approaches, this model incorporates richer prior knowledge, transforming the previous black-box model into a more interpretable framework. This enhanced interpretability and knowledge integration allow the model to be more effectively applied in practical scenarios. To validate the performance of the model designed in this study, we set up a smart home scenario for consumer electronics. In this scenario, when a consumer electronic device malfunctions, such as a rupture within an air conditioner's internal piping, users may find it challenging to detect the issue. At this point, the user uploads fault information (e.g., fault description, data collected from sensors). The model then extracts multimodal features through the feature extraction network, which are input into the GraphSAGE graph neural network for reasoning. The inference reveals an excessive level of CF_2Cl_2, which is then matched with fault nodes in the knowledge graph. Using a semantic matching template, the model identifies the fault entity (air conditioner) and provides a diagnosis result (Freon leakage, the internal pipes of the air conditioner are damaged), as illustrated in Figure 6.

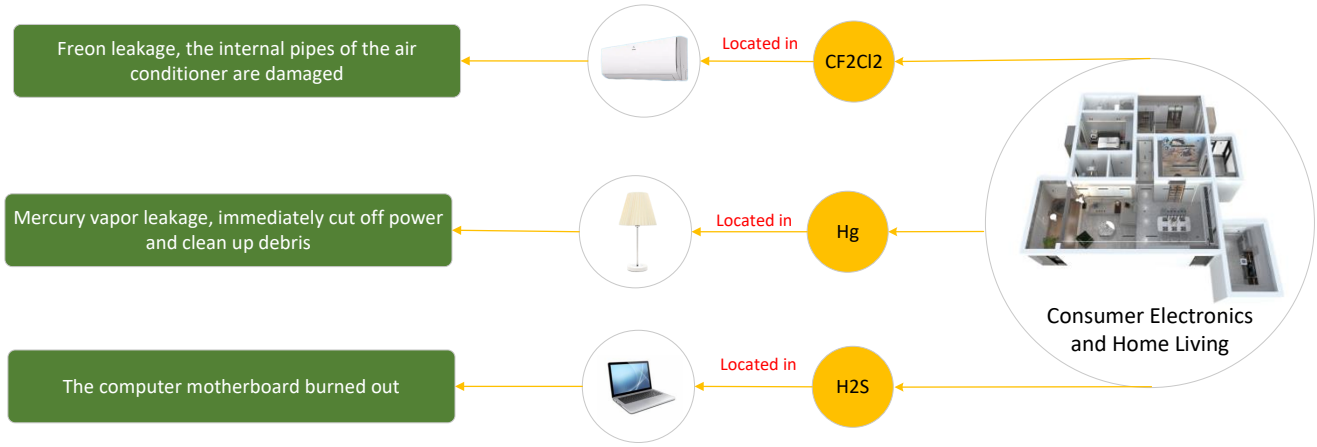


Figure 6. Knowledge graph reasoning.

Using GraphSAGE to process multimodal data enhances the model's understanding and handling of fault data, ensuring that the resulting fault diagnosis outcomes are interpretable. Table 4 shows the inference results of the multimodal knowledge graph in consumer electronics and home scenarios. The fault results obtained through GraphSAGE are matched with knowledge graph entities to identify the head entity. Subsequently, the knowledge graph inference retrieves corresponding fault entities such as "air conditioner", "smart bulb", and

"computer". Once the relevant fault entities are identified, the appropriate inference results are derived through entity relationship matching within the knowledge graph.

Table 4. Inference result.

Knowledge Graph Entity Matching	Fault Entities	The Inference Results from the Knowledge Graph
CF2Cl2 (81%)	Air Conditioner	Freon leakage, the internal pipes of the air conditioner are damaged.
Hg (69%)	Smart Bulb	Mercury vapor leakage, immediately cut off power and clean up debris
H2S (61%)	Computer	The computer motherboard burned out

The inference results from the knowledge graph are input into the text generation model to produce fault diagnosis text. During the text generation phase, to evaluate the effectiveness of the consumer electronics fault text generation model, the fine-tuned CPT model is trained and compared with other advanced models on the same dataset. The experimental results are shown in Table 5.

Table 5. Comparison of text generation model performance.

Models	LCSTS (Rouge-L)	CSL (Rouge-L)	ADGEN (BLEU-4)
mT5	33.5	56.7	10.2
BART	37.8	62.1	9.9
RoBERTa	41.0	-	-
Ours	42.0	63.7	10.7

In the consumer electronics fault text generation task, this model is compared with the mT5 model, showing improvements of 8.5 and 7 points in LCSTS (Rouge-L) and CSL (Rouge-L), respectively, and a 0.5-point increase in ADGEN (BLEU-4). Compared to the BART model, LCSTS (Rouge-L) and CSL (Rouge-L) improved by 4.2 and 1.6 points, respectively, and ADGEN (BLEU-4) increased by 0.8 points. When compared with the RoBERTa model, LCSTS (Rouge-L) improved by 1 point. This is because the model in this study improves text generation performance through additional pre-training on natural language generation tasks. These results indicate that our model outperforms others in the consumer electronics fault text generation task.

To further demonstrate the superiority of our model, we conducted comparative experiments with state-of-the-art models in the field, using the same dataset across all models. We selected the CLIP model for comparison. CLIP is a model that unifies image and text embeddings into a shared space, using contrastive learning to train image and text representations simultaneously. This makes CLIP a suitable baseline for evaluating multimodal data processing performance. Additionally, we chose GCN and GAT models for comparison, as both are deep learning models designed for handling graph-structured data and are widely applied in tasks such as node classification and edge prediction across various domains, including image, text, and social network analysis. Therefore, GCN and GAT serve as benchmarks for evaluating performance on complex graph structures. We selected accuracy, precision, and F1 score as evaluation metrics. Each model was tested multiple times, and the average of these experimental results was taken as the final outcome. The experimental results are shown in Table 6.

Table 6. Comparison of experimental results.

Models	Accuracy (%)	Precision (%)	F1 Score (%)
GCN	82.2	82.1	72.3
GAT	83.5	83.5	74.8
CLIP	88.1	86.3	78.9
Ours	92.1	89.4	80.8

The analysis of the experimental results shows that our model achieved excellent performance across all metrics, indicating that it outperforms other models in the task of consumer electronics fault diagnosis. The performance of CLIP ranks second, attributed to its significant advantage in multimodal feature alignment, allowing it to achieve effective multimodal fusion. However, its performance in handling complex graph structures is inferior to our model. GCN and GAT demonstrated moderate performance, as their methods for aggregating neighbor features are somewhat limited, which significantly affects their performance. The experimental results indicate that, compared to other models, our model demonstrates superior performance.

4.6. Ablation Study

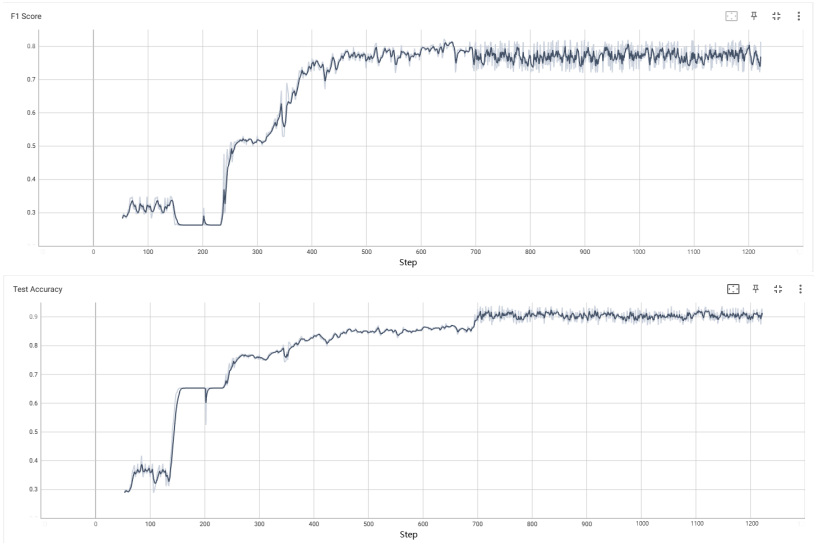
To demonstrate the rationale and effectiveness of our proposed method, we designed an ablation study. For the four input modalities of the model—image, text, audio, and olfactory information—we conducted the following experiments: removing the image modality while retaining text, audio, and olfactory modalities; removing the text modality while retaining image, audio, and olfactory modalities; removing the olfactory modality while retaining image, text, and audio modalities; and removing the audio modality while retaining image, text, and olfactory modalities. In each experiment, we kept the rest of the model unchanged and used the same training and testing sets to ensure the comparability of the experimental results. The experimental results are shown in Figure 7.

To quantify the contribution of each modality, we used ACC and F1 as evaluation metrics. In each ablation experiment, we compared the model's performance changes in these metrics to assess the impact of removing a particular modality on the model's performance. The experimental results are shown in Table 7.

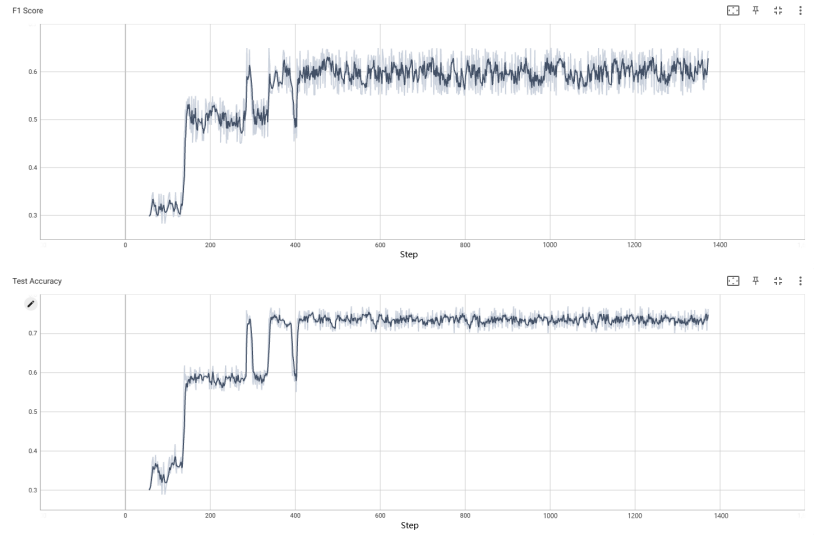
Table 7. The impact of different modalities on model performance.

Ablation Experimental Conditions	F1 (%)	Accuracy (%)
Ours	80.8	92.1
Image Modality Removed	63.2	75.5
Text Modality Removed	53.7	60.0
Smell Modality Removed	77.7	85.7
Audio Modality Removed	76.5	84.8

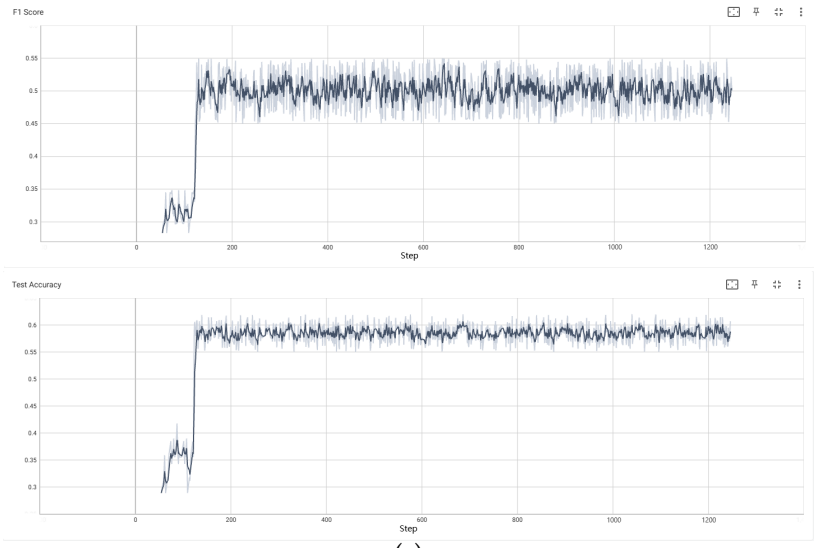
Based on the results of the ablation experiments, we can observe that the model's performance declines when certain modalities are removed. Analyzing the experimental results, we can see that when the text modality is removed, both the F1 score and accuracy of the model drop significantly, indicating that fault text information plays a crucial role in fault description and diagnosis. Similarly, the removal of the image modality leads to a noticeable performance decrease, suggesting that image fault information also has a significant impact on diagnostic performance. In contrast, the audio and olfactory modalities have a relatively smaller impact on the model's final decision. While these modalities can provide useful information in specific scenarios, their effect on overall model performance is less pronounced than that of the image and text modalities. The results of the ablation experiments validate the rationale behind the proposed method.



(**a**)

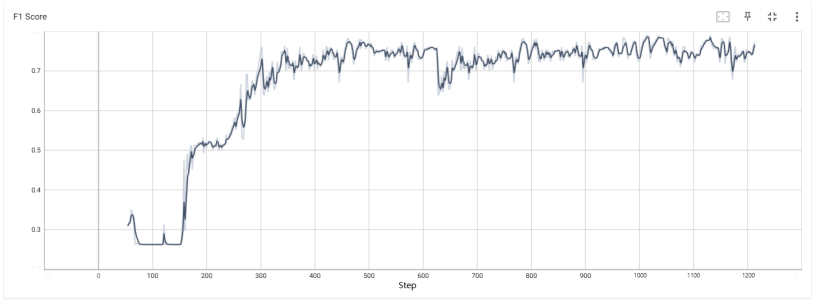


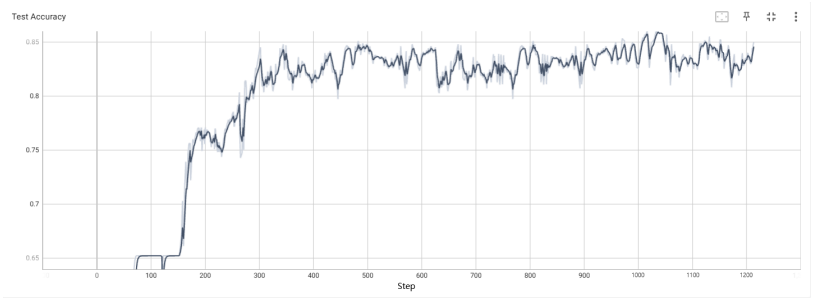
(**b**)



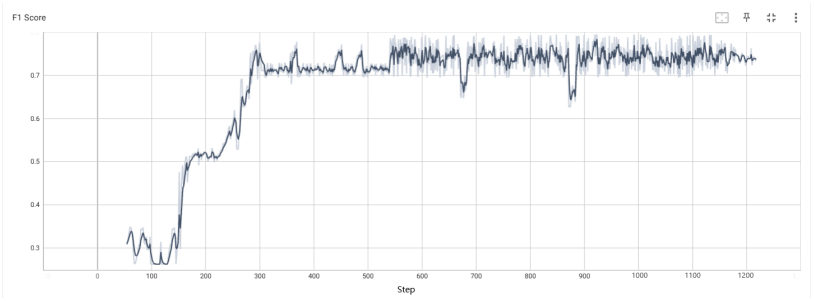
(**c**)

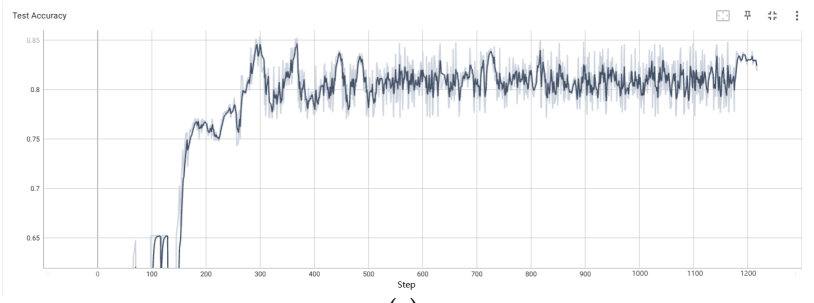
Figure 7. *Cont*.





(**d**)





(**e**)

Figure 7. Ablation experiment results. (**a**) Ours. (**b**) Image modality removed. (**c**) Text modality removed. (**d**) Smell modality removed. (**e**) Audio modality removed.

4.7. Assessment of the BitFit Fine-Tuning Method

To enhance the CPT language model's suitability for the task of consumer electronics fault text generation, we fine-tuned the model. Various tasks require different fine-tuning methods; based on the characteristics of the consumer electronics fault text generation task, we chose the BitFit method. BitFit is a sparse fine-tuning approach that only adjusts the bias terms of the model. For small- to medium-sized training datasets, applying BitFit to the model is more efficient than fine-tuning the entire model. Moreover, for larger datasets, this method has advantages over other sparse fine-tuning techniques. Utilizing the BitFit method to fine-tune the CPT model allows for the full utilization of the rich information from the multimodal knowledge graph, while ensuring model efficiency and generalization, reducing computational costs for training and deployment, and ultimately improving the overall performance of fault diagnosis in consumer electronics.

To validate the advantage of the proposed method in the consumer electronics fault text generation task, we trained and compared the original CPT model with the CPT model fine-tuned using the BitFit method on the same dataset and under identical experimental conditions. Rouge-1, Rouge-2, and Rouge-L were used as evaluation metrics. The comparison results are shown in Figure 8, where Figure 8a–d represents the original CPT model, and Figure 8e–h represents the BitFit fine-tuned model. By comparing Figure 8a–c,e–g, it is observable that the BitFit fine-tuned CPT model achieves higher scores across multiple evaluation metrics, with notable improvements in Rouge-1, Rouge-2, and Rouge-L metrics. Additionally, comparing Figure 8d,h shows that the BitFit fine-tuned CPT model has a lower loss function and faster convergence rate, underscoring the effectiveness of the BitFit fine-tuning approach.

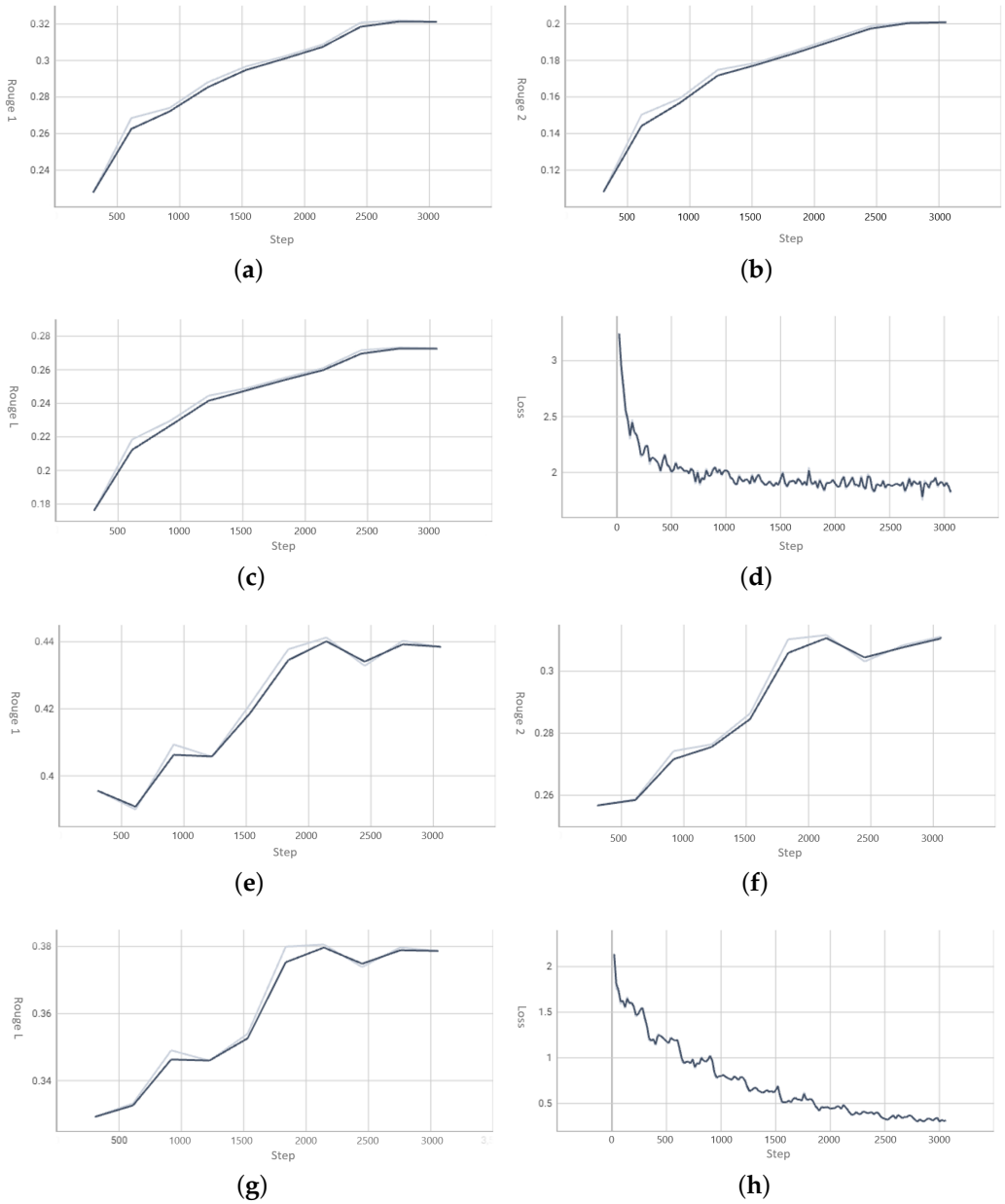


Figure 8. Performance comparison of the model fine-tuned with BitFit.

4.8. Case Evaluation

The model utilizes multimodal information uploaded by users during consumer electronics failures as diagnostic samples. The richness and complementarity of these data

ensure the accuracy of the generated fault diagnosis text. Figure 9 shows a specific case of fault diagnosis using a multimodal knowledge graph-based text generation method. The user provides an image of the product fault, along with a voice or text description and odor information. The model uses this multimodal information to perform a fault diagnosis and generate results.

In the presented case, the diagnostic result is as follows: 'It was found that the battery had severely deteriorated. The battery needs to be replaced'. The actual fault text is as follows: 'It was discovered that the issue was caused by battery aging. The battery needs to be replaced'. By comparing the model-generated result with the actual text, a text generation accuracy of 98.4% is calculated. Additionally, based on experimental results, the proposed method shows strong performance in diagnosing faults in other consumer electronics, achieving an accuracy rate of 97.3%.

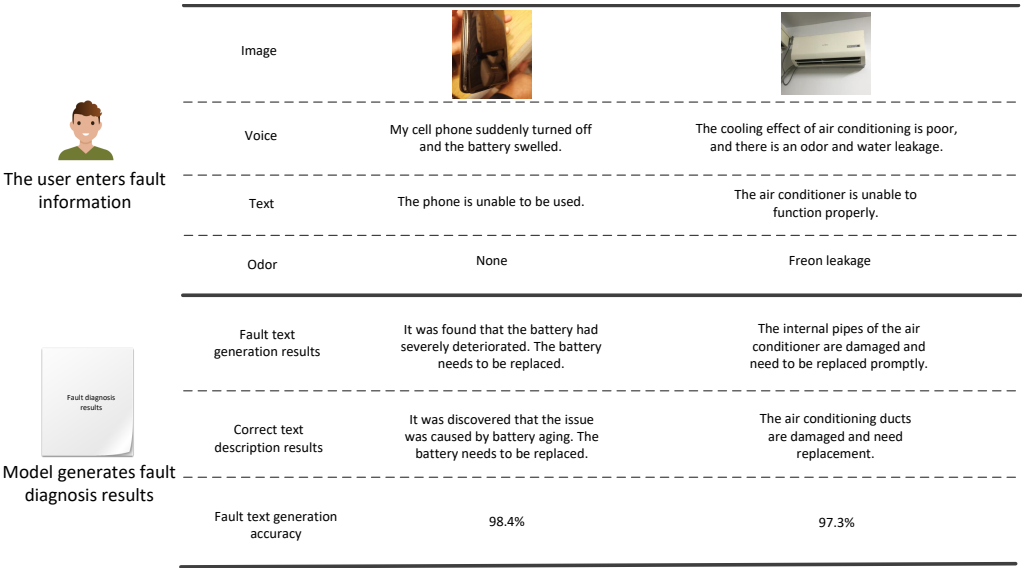


Figure 9. Specific cases of consumer electronics fault diagnosis.

5. Conclusions

The focus of this research is fault diagnosis technology for consumer electronics. Unlike existing methods that rely on single-modality fault diagnosis or solely use knowledge graphs, this approach combines multimodal information with knowledge graphs, enabling fault reasoning from multiple perspectives and ensuring reliable diagnosis results.

Specifically, the process begins by using a graph neural network to handle the multimodal fault information provided by the user, fully utilizing the data to generate a fault diagnosis vector. This diagnosis vector is then matched with knowledge graph entity vectors based on similarity, allowing inference of fault points and corresponding solutions. Finally, an advanced text generation model transforms the inferred results into interpretable fault descriptions, presenting the diagnostic outcome to the user. This makes the approach more applicable to real-world scenarios.

6. Future Work

In future work, we plan to incorporate advanced techniques to enhance the model's performance. This includes introducing an attention mechanism into the graph convolutional network to improve the model's ability to distinguish the importance of different modal features. We also intend to integrate the transformer module to enhance text feature extraction, ensuring effective capture of text features even in complex scenarios. Additionally, we aim to use advanced NLP models within the multimodal knowledge graph to further improve text feature quality. With advancements in sensor technology, we plan to utilize high-precision, next-generation sensors to increase the accuracy of fault detection.

In the future, we plan to expand the application scenarios of our model to other domains. For instance, in industrial equipment maintenance, the model could integrate sensor data, image data, and operator descriptions to facilitate maintenance. We also aim to apply the model to real-time fault detection in smart home products, allowing it to integrate multimodal data and detect potential anomalies in real time. Additionally, we intend to extend its application to medical device fault diagnosis and IoT network security [37]. In medical device diagnosis, integrating multimodal data (e.g., signal data and operation records) can improve fault detection reliability. In IoT network security, fusing data from various devices could help detect potential security threats or anomalies at an earlier stage.

Although the proposed model has certain advantages, it also has limitations, such as difficulties in handling missing or inaccurate data. To address these challenges, we plan to introduce generative model-based data imputation techniques to enhance performance. Additionally, to mitigate potential noise, we are considering integrating adversarial training during feature extraction to improve the model's stability in dealing with noisy data.

Author Contributions: Software, Q.L.; Validation, X.Z.; Writing—original draft, Y.S.; Writing—review and editing, Y.W. and Y.S.; Visualization, Y.S.; Supervision, L.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China No. 62106074, in part by the Scientific Research Fund of Hunan Provincial Education Department under Grant No. 22A0422, No. 23C0194 and in part by the Hunan Provincial Innovation Foundation for Postgraduate under Grant No. CX20220835.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in [github] at [https://github.com/fastnlp/CPT].

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

References

1. Iannacci, J. Internet of things (IoT); internet of everything (IoE); tactile internet; 5G—A (not so evanescent) unifying vision empowered by EH-MEMS (energy harvesting MEMS) and RF-MEMS (radio frequency MEMS). *Sens. Actuators A Phys.* **2018**, *272*, 187–198. [CrossRef]
2. Cao, K.; Liu, Y.; Meng, G.; Sun, Q. An overview on edge computing research. *IEEE Access* **2020**, *8*, 85714–85728. [CrossRef]
3. Farias da Costa, V.C.; Oliveira, L.; de Souza, J. Internet of everything (IoE) taxonomies: A survey and a novel knowledge-based taxonomy. *Sensors* **2021**, *21*, 568. [CrossRef] [PubMed]
4. Bhargava, C.; Banga, V.K.; Singh, Y. Failure prediction and health prognostics of electronic components: A review. In Proceedings of the 2014 Recent Advances in Engineering and Computational Sciences (RAECS), Chandigarh, India, 6–8 March 2014; pp. 1–5.
5. Miraz, M.H.; Ali, M.; Excell, P.S.; Picking, R. A review on Internet of Things (IoT), Internet of everything (IoE) and Internet of nano things (IoNT). In Proceedings of the 2015 Internet Technologies and Applications (ITA), Wrexham, UK, 8–11 September 2015; pp. 219–224.
6. Pecht, M.; Tuchband, B.; Vichare, N.; Ying, Q.J. Prognostics and health monitoring of electronics. In Proceedings of the 2007 International Conference on Thermal, Mechanical and Multi-Physics Simulation Experiments in Microelectronics and Micro-Systems. EuroSime 2007, London, UK, 16–18 April 2007; pp. 1–8.
7. Kullu, O.; Cinar, E. A Deep-Learning-Based Multi-Modal Sensor Fusion Approach for Detection of Equipment Faults. *Machines* **2022**, *10*, 1105. [CrossRef]
8. Nan, X.; Zhang, B.; Liu, C.; Gui, Z.; Yin, X. Multi-Modal Learning-Based Equipment Fault Prediction in the Internet of Things. *Sensors* **2022**, *22*, 6722. [CrossRef]
9. Yang, Q.; Li, H.; Hua, C.; Wu, D. Fault Diagnosis of Multimodal Chemical Process Based on Attention Mechanism and Deep Subdomain Countermeasure Adaptive Network. In Proceedings of the 2022 2nd International Conference on Algorithms, High Performance Computing and Artificial Intelligence (AHPCAI), Guangzhou, China, 21–23 October 2022; pp. 686–690.
10. Pecht, M.; Gu, J. Physics-of-failure-based prognostics for electronic products. *Trans. Inst. Meas. Control* **2009**, *31*, 309–322. [CrossRef]
11. Kumar, S.; Torres, M.; Chan, Y.; Pecht, M. A hybrid prognostics methodology for electronic products. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–8 June 2008; pp. 3479–3485.

12. Kumar, S.; Chow, T.W.; Pecht, M. Approach to fault identification for electronic products using Mahalanobis distance. *IEEE Trans. Instrum. Meas.* **2009**, *59*, 2055–2064. [CrossRef]
13. Stetter, R.; Phleps, U. Design for diagnosis. In Proceedings of the DS 68-5: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society Through Engineering Design, Vol. 5: Design for X/Design to X, Lyngby/Copenhagen, Denmark, 15–19 August 2011.
14. Cai, B.; Liu, Y.; Xie, M. A dynamic-Bayesian-network-based fault diagnosis methodology considering transient and intermittent faults. *IEEE Trans. Autom. Sci. Eng.* **2016**, *14*, 276–285. [CrossRef]
15. Nguyen, D.T.; Duong, Q.B.; Zamai, E.; Shahzad, M.K. Fault diagnosis for the complex manufacturing system. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2016**, *230*, 178–194. [CrossRef]
16. Arcos, B.P.; Bakker, C.; Flipsen, B.; Balkenende, R. Practices of fault diagnosis in household appliances: Insights for design. *J. Clean. Prod.* **2020**, *265*, 121812. [CrossRef]
17. Lin, W. Intelligent Fault Diagnosis of Consumer Electronics Sensor in IoE via Transformer. *IEEE Trans. Consum. Electron.* **2023**, *70*, 1259–1267. [CrossRef]
18. Jin, X.; Zhao, M.; Chow, T.W.; Pecht, M. Motor bearing fault diagnosis using trace ratio linear discriminant analysis. *IEEE Trans. Ind. Electron.* **2013**, *61*, 2441–2451. [CrossRef]
19. Souri, A.; Mohammed, A.S.; Potrus, M.Y.; Malik, M.H.; Safara, F.; Hosseinzadeh, M. Formal verification of a hybrid machine learning-based fault prediction model in Internet of Things applications. *IEEE Access* **2020**, *8*, 23863–23874. [CrossRef]
20. Mao, Y.; Bhuse, V.; Zhou, Z.; Pichappan, P.; Abdel-Aty, M.; Hayafuji, Y. Applied mathematics and algorithms for cloud computing and iot. *Math. Probl. Eng.* **2014**, *2014*, 946860. [CrossRef]
21. Fernandes, S.; Antunes, M.; Santiago, A.R.; Barraca, J.P.; Gomes, D.; Aguiar, R.L. Forecasting appliances failures: A machine-learning approach to predictive maintenance. *Information* **2020**, *11*, 208. [CrossRef]
22. Song, K. Research on Intelligent Fault Diagnosis Technology of Power Electronic Circuits. In Proceedings of the 2021 International Conference on Control Science and Electric Power Systems (CSEPS), Shanghai, China, 28–30 May 2021; pp. 291–294.
23. Lall, P.; Gupta, P.; Angral, A. Anomaly detection and classification for PHM of electronics subjected to shock and vibration. *IEEE Trans. Compon. Packag. Manuf. Technol.* **2012**, *2*, 1902–1918. [CrossRef]
24. Marijan, D.; Zlokolica, V.; Teslic, N.; Pekovic, V.; Tekcan, T. Automatic functional TV set failure detection system. *IEEE Trans. Consum. Electron.* **2010**, *56*, 125–133. [CrossRef]
25. Li, X.; Gao, J.; Li, W. Reliability Design Method of Digital Diagnosis and Treatment Equipment. In *Reliability Engineering Technology of Digital Diagnosis and Treatment Equipment*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 35–138.
26. Liang, X.; Zhao, Y.; Zhou, Z. Research on application of fuzzy fault tree analysis in the electronic equipment fault diagnosis. In Proceedings of the 2nd International Conference on Computer and Automation Engineering (ICCAE), Singapore, 26–28 February 2010; Volume 2, pp. 65–67.
27. Wu, Y.; Liu, F.; Wan, L.; Wang, Z. Intelligent Fault Diagnostic Model for Industrial Equipment Based on Multimodal Knowledge Graph. *IEEE Sens. J.* **2023**, *23*, 26269–26278. [CrossRef]
28. Zhang, B.; Liu, X.; Ji, S.; Yang, X.; Xie, T.; Yu, Y.; Li, F. Research and application of power transformer technical standard retrieval algorithm based on knowledge graph. In Proceedings of the 2022 4th International Academic Exchange Conference on Science and Technology Innovation (IAECST), Guangzhou, China, 10–12 December 2022; pp. 30–34.
29. Deng, J.; Wang, T.; Wang, Z.; Zhou, J.; Cheng, L. Research on event logic knowledge graph construction method of robot transmission system fault diagnosis. *IEEE Access* **2022**, *10*, 17656–17673. [CrossRef]
30. Luo, R.; Sun, L.; Xia, Y.; Qin, T.; Zhang, S.; Poon, H.; Liu, T.Y. BioGPT: Generative pre-trained transformer for biomedical text generation and mining. *Brief. Bioinform.* **2022**, *23*, bbac409. [CrossRef]
31. Lebret, R.; Grangier, D.; Auli, M. Neural text generation from structured data with application to the biography domain. *arXiv* **2016**, arXiv:1603.07771.
32. Ji, Z.; Lee, N.; Frieske, R.; Yu, T.; Su, D.; Xu, Y.; Ishii, E.; Bang, Y.J.; Madotto, A.; Fung, P. Survey of hallucination in natural language generation. *ACM Comput. Surv.* **2023**, *55*, 1–38. [CrossRef]
33. Xu, P.; Patwary, M.; Shoeybi, M.; Puri, R.; Fung, P.; Anandkumar, A.; Catanzaro, B. MEGATRON-CNTRL: Controllable story generation with external knowledge using large-scale language models. *arXiv* **2020**, arXiv:2010.00840.
34. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. Learning transferable visual models from natural language supervision. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; pp. 8748–8763.
35. Shao, Y.; Geng, Z.; Liu, Y.; Dai, J.; Yan, H.; Yang, F.; Zhe, L.; Bao, H.; Qiu, X. Cpt: A pre-trained unbalanced transformer for both chinese language understanding and generation. *arXiv* **2021**, arXiv:2109.05729. [CrossRef]
36. Geier, M.J. *How to Diagnose and Fix Everything Electronic*; McGraw-Hill/TAB Electronics: New York, NY, USA, 2011.
37. Bhuva, D.R.; Kumar, S. A novel continuous authentication method using biometrics for IOT devices. *Internet Things* **2023**, *24*, 100927. [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.