Mathematical Biosciences and Engineering

MBE, 20 (6): 9992–10006.
DOI: 10.3934/mbe.2023438
Received: 31 December 2022
Revised: 28 February 2023
Accepted: 02 March 2023
Published: 27 March 2023

http://www.aimspress.com/journal/MBE

Research article

An intelligent fault detection approach for digital integrated circuits through graph neural networks

Zulin Xu*

SDU-ANU Joint Science College, Shandong University, Shandong 264209, China

* **Correspondence:** Email: x15965196954@126.com.

Abstract: To quickly and accurately realize the fault diagnosis of analog circuits, this paper introduces the graph neural network method and proposes a fault diagnosis method for digital integrated circuits. The method filters the signals present in the digital integrated circuit to remove noise signals and redundant signals and analyzes the digital integrated circuit characteristics after the filtering process to obtain the digital integrated circuit leakage current variation. To the problem of the lack of a parametric model for Through-Silicon Via (TSV) defect modeling, the method of TSV defect modeling based on finite element analysis is proposed. The common TSV defects such as voids, open circuits, leakage, and unaligned micro-pads are modeled and analyzed by using industrial-grade FEA tools Q3D and HFSS, and the equivalent circuit model of resistance inductance conductance capacitance (RLGC) for each defect is obtained. Finally, the superior performance of this paper in fault diagnosis accuracy and fault diagnosis efficiency is verified by comparing and analyzing with the traditional graph neural network method and random graph neural network method for active filter circuits.

Keywords: graph neural network; intelligent fault detection; digital integrated circuits; artificial intelligence

1. Introduction

With the development of the national economy, people's demand for power supply reliability is increasing, and the distribution network is an important link between the power system and users, so when the distribution feeder fault occurs, whether the fault line can be found in time, and timely restoration of power supply is an important guarantee to improve the reliability of power supply [1]. According to statistics, the distribution network line fault single-phase grounding fault rate accounted

for about 80% of the total fault, the advantage of a small-current grounding system is a single-phase grounding fault, because the grounding phase current is not large so that the line between the three phases of the electrical skin remains symmetrical, the load power supply temporarily does not affect. The system can continue to run for 1 to 2 hours, without immediately removing the grounding phase, the circuit breaker does not have to immediately jump Min [2]. This ensures an uninterrupted continuous power supply to users and improves the reliability of the power supply. However, as the scale of the distribution network continues to grow, the line continues to increase, and the number of cable lines and Weilian mixed lines is also increasing, when a single-phase fault occurs, the grounding capacitance current also increases, a long time with fault operation, so that the arc grounding caused by the system overvoltage is too high. The arc is difficult to self-extinguish if the fault cannot be removed in time, will damage equipment, heavy power plant unit shutdown, process interruption, and other malignant accidents, and damage the safe operation of the system.

As the feature size of transistors continues to shrink, the development of integrated circuits is becoming a bottleneck. In the face of the demand for IC development and innovation, Three-dimensional Integrated Circuit (3D IC) based on Through-Silicon Via (TSV) interconnects multilayer wafers vertically through the TSV structure to achieve higher integration, smaller size, lower latency, and power consumption [3]. It is the trend of next-generation ICs with higher integration, smaller size, lower latency, and power consumption. In 3D ICs, TSV is a new type of interconnect structure with an immature process, which is prone to multiple failures during production and wafer binding. Testing and troubleshooting TSVs during the Wafer Probe stage can ensure the validity and reliability of TSVs, improve the yield of 3D ICs, and reduce manufacturing costs. Due to the technical difficulty of pre-binding TSV testing, it is difficult to obtain sufficient fault information for troubleshooting at this stage, so pre-binding TSV testing is mostly focused on detecting the presence or absence of faults [4]. The post-binding TSV test, however, has more test resources at its disposal and more relaxed test access restrictions, so the post-binding TSV test can be used not only to detect TSV defects generated during the binding process but also to diagnose faults by obtaining more abundant fault information. Failure analysis is as comprehensive and accurate as possible, thus providing valuable information for weaknesses in design and manufacturing solutions, improving design and process parameters, and contributing to higher yield.

Integrated is through the capacitors, resistors, transistors, and other electronic devices composed of the existence of a specific function, independent of the overall circuit. The causes of failure of ICs in different life periods are different, ICs can be divided into three stages: early death, normal cycle, and aging period [5]. The cause of IC failure in the normal cycle is uncertain, most of them are caused by indirect failure, IC failure in the normal cycle is not permanent, and usually can be repaired, and IC failure rate in the normal cycle has stability. Increasing service life is the main cause of IC failure during the aging period, as the IC is affected by the workload, stress time, and other factors, resulting in the degradation of the IC performance and aging phenomenon. With the rapid progress of integrated circuit design and manufacturing technology, the number of transistors integrated on a single chip has tens of billions, and the complexity and scale of integrated circuits are greatly increased [6]. These problems directly lead to an increase in test cost, so this paper proposes a digital IC fault detection method based on K-means clustering and graph neural network, which has important theoretical and practical value for the research of IC test methods.

2. Related work

Fault diagnosis is a key technology to ensure the reliable and efficient continuous operation of industrial systems. With the in-depth development of industrial informatization and intelligence, fault diagnosis methods gradually move from signal processing-based methods to deep learning-based methods [7]. In recent years, scholars have applied wavelet analysis, the Prony method, artificial neural network, information fusion, fuzzy method, etc. to small current grounding protection. Although the problem of difficult extraction of fault features has been solved to some extent, the analysis of the fault features themselves has been neglected. The graph neural network (GNN) is a class of methods developed based on deep learning to process graph data, which has better performance and interpretability. GNN can establish the correlation between nodes in the graph and is increasingly used in the fields of knowledge graphs, social networks, etc. GNN represents all the deep learning methods used to study the characteristics of graph data, breaking through the traditional deep learning methods are difficult to GNN has better performance and interpretability, and is mainly applied to data with non-Euclidean domains, using the powerful ability of graph to represent data to solve the problems that are difficult to solve in regular space [8]. To address the problem of low fault resolution of traditional testability models in analog circuit diagnosis, Balouchestani M et al. proposed a fault pair Boolean table-based testability analysis and fault diagnosis method for analog circuits [9]. Deng A et al. proposed Cluster-GCN by sampling subgraphs using a graph clustering algorithm and performing neighborhood search and graph convolution operations in the sampled subgraphs, thus making Cluster-GCN not only able to explore larger graphs but also use deeper graph structures with less time and memory [10].

The k-means clustering algorithm is an unsupervised learning algorithm based on division, which is a representative algorithm of cluster analysis. In the current era of big data, the use of the K-means algorithm to process data is also becoming more frequent, and the drawbacks of the K-means clustering algorithm are gradually highlighted: the need to determine k clusters before the algorithm runs, sensitivity to the initial centroid selection, and high influence by outliers, etc.

Being in the era of big data, we are always accompanied by challenges, but while the times bring us many challenges, they also create more opportunities for us, and it is in this complex environment that data mining technology is continuously developed and improved [11]. The K-means algorithm, which is one of the most classic algorithms for classifying clusters, has become one of the most popular algorithms in reality because of its simplicity and effectiveness. Today, the K-means algorithm is still widely used due to its excellent performance, especially when dealing with data with numerical attributes, due to the simplicity and fast convergence of the algorithm, which is more geometrically and statistically advantageous [12–15]. From the perspective of GNN, the relationship between feature data points is explored, signal features are extracted, and the problem of missing and mixed composite fault signals is solved by analyzing composite faults that have not yet appeared and may appear from known fault signals. The K-means clustering analysis is combined with the validity index analysis, and the clustering center is selected as the pattern layer neuron to train the probabilistic neural network model, thus reducing the complexity of the model and greatly reducing the fault diagnosis time [16]. Meanwhile, to address the problem that real digital IC faults are difficult to obtain and real fault data are scarce, features are learned from artificially simulated damaged fault data and migrated to real in-service faults to construct a comprehensive digital IC fault diagnosis model for diagnosing single faults and compound faults, and to realize migration diagnosis from artificially simulated damaged faults to

in-service faults.

3. Detection method based on K-means clustering and graph neural network

3.1. Fault routing method based on k-means clustering analysis

The K-means algorithm is a relatively simple unsupervised learning method, first proposed by Macqueen in 1867, which requires a given number of samples to be classified, K, and then divides the realized input N data into K clusters according to certain rules, so that the objects in the obtained clusters are highly similar and the objects in different clusters are less similar [17]. The core idea of this algorithm is to cluster the K points in the space as the center, group the objects closest to them into one class, and gradually update the respective cluster centers through continuous iterative computation.

First, Num initial cluster centers are randomly selected from the data set; then the remaining data points are assigned to the clusters of these Num points according to the minimum distance criterion; next, the centers of each cluster are recalculated and the clusters are updated until the criterion function converges.

$$J = \sum_{i=1}^{Num} \left| p - m_i \right|^2, p \in c_i . \tag{1}$$

where J is the error sum of squares of all data objects in a given data set; p is the data points contained in the cluster c_i and m_i is the center of mass of the cluster c_i. The criterion tries to make the Num clusters as compact as possible between the intra-cluster elements and as far away as possible between the inter-cluster elements.

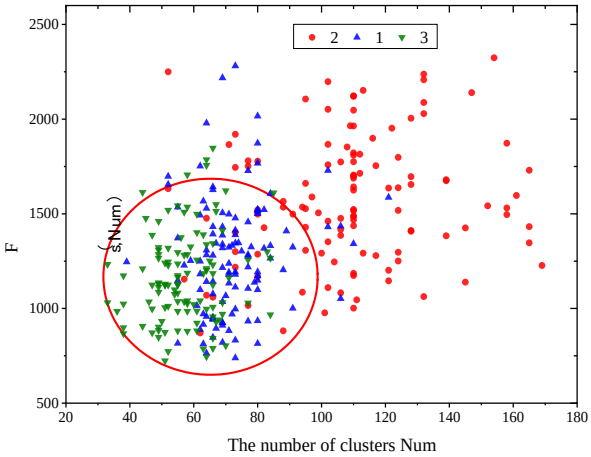

Figure 1. Determination of the number of clusters.

When finding the clustering centers, since the initial clustering centers of the traditional K-means algorithm are chosen randomly, the resulting clustering results will have great volatility when the clustering centers chosen by the K-means algorithm are different, i.e., it is not appropriate to choose the initial clustering centers randomly [18]. Therefore, this chapter gives a new initial clustering center

selection algorithm based on the principle that the clusters formed by the clustering centers should make the clusters have a high density within a small threshold radius, and the clustering centers are far from each other, as shown in Figure 1.

The db10 wavelet is used to decompose the transient zero-sequence current obtained from the Matlab electromagnetic transient simulation within 5ms after the fault to obtain the wavelet coefficients, and the energy is calculated from the square of the wavelet coefficients after the single branch reconstruction, and then the total energy of the transient zero-sequence current in the full frequency band is calculated [19]. At the same time, the integrated wavelet energy relative fullness is calculated, and the two dimensions of transient zero-sequence current energy and integrated wavelet energy relative fullness are used as measures to characterize the fault characteristics and are mapped to the two-dimensional plane [20]. The k-means cluster analysis algorithm is used to calculate the clustering centers of the data. One class is the faulty line and the other class is the sound line, and the clustering centers of the two classes are noted separately. The faulty lines are determined based on the distance between the test data and the clustering centers of the two classes.

From the theoretical analysis and simulation, analysis can be seen, when the line fault occurs, due to the existence of a large current flow channel of the fault line, the fault line measurement end of the fault current detected in a period window of the general trend for the rising, while the non-fault line measurement end of the fault current component detected in a period of the time window, its blocked current component is generally down. There are significant differences in waveforms between faulty and non-faulty lines. The k-means clustering analysis is used to describe and characterize the current waveforms observed at the measurement end of the faulty and sound lines, and the K-means clustering analysis-based fault line selection criteria for the distribution network are constructed.

In analog circuits, when different faults occur, the faults behave differently, thus causing different test signal fault information to be collected at different test nodes. For different fault classes, if the intra-class distance of the same fault class is small, and the inter-class distance of different fault classes is large, then the fault is more distinguishable at that time [21]. And under different fault conditions, if the inter-class distance of the smallest different fault class is larger and the intra-class distance of the largest same fault class is smaller, then the information collected at this test point is more distinguishable, so this test point is treated as a sensitive test node.

The average intra-class distance of the collected class j fault samples at the *i*-th test node is defined as

$$d_{i,j} = \frac{1}{N} \sum_{a,b=1} \left| f(\mathrm{a}_i) - f(\mathrm{b}_j) \right|. \tag{2}$$

The average of the intra-class distances for *Y* fault classes at the *i*-th test node D_i is defined as

$$D_i = \sum_{i=1}^{Y} d_{i,j}. \tag{3}$$

The sensitivity factor of the test node λ_i indicates the ability of the *i*th test node to classify different fault categories. The larger the value λ_i, the more sensitive the *i*th test node is to the characteristics of *Y* faults, and therefore the stronger the ability to classify different faults, which is beneficial to the subsequent diagnosis. The fault sensitivity factor of different test nodes is analyzed, and multiple test nodes of the circuit under test are sorted according to the value of the fault sensitivity

factor from the largest to the smallest, to ensure that the test nodes with stronger fault classification recognition ability can enter the selection range of the optimal test nodes in priority and avoid the blindness of test node selection.

3.2. Fault detection model design for digital integrated circuits based on K-means clustering and graph neural networks

Convolutional Neural Networks (CNNs) are good at processing canonical matrices, but not all data can be integrated into a standard matrix form, and CNNs cannot extract and learn features when the data in cases such as social networks are integrated into unordered graph structures. To solve this problem, the Graph Convolutional Network (GCN) was developed to extend the processing capability of non-Euclidean distance data by introducing graph structure, and providing a method to extract features from irregular data, so it is widely used in business recommendation systems, road forecasting, financial risk control, etc. Since graph structure can represent the connection between data and data through nodes and edges, GCN can also be used in the field of fault diagnosis to solve the problem of insufficient training due to too few labeled samples [22]. In addition to the graph structure, GCN can obtain excellent output results even with the initial parameters of W_l without optimization training, so GCN can achieve excellent output results with only 2–3 layers without multi-layer superposition.

Spectral-domain convolution mainly relies on graph theory to realize convolution operations on irregular graphs and explore the properties of graphs. The pooling operation aims to reduce the size of parameters by generating smaller feature representations by down-sampling feature points, thus avoiding overfitting, solving the problems of substitution invariance and computational complexity, and is an essential operation to reduce the size of graphs [23]. Mean pooling and average pooling are the most common and effective methods to implement downsampling, because it is fast to calculate the mean or maximum value in the merge window, and it is necessary to perform simple pooling operations during network training to reduce the dimensionality of the graph data and mitigate the cost of graph transformation operations [24].

To reduce the computational effort without affecting the convolution effect, a local first-order approximation is used to limit the convolution operation, i.e., let, the concise equation be as follows.

$$G(\mathrm{x}) = \varepsilon(\mathrm{I}_n - \frac{D^{-1/2}}{AD^{1/2}}). \tag{4}$$

To prevent gradient extinction or gradient explosion, renormalization is required. The propagation rule for the convolution layer in GCN is expressed as

$$H_{l+1} = \delta(\frac{D^{-1/2}}{AD^{1/2}} + \frac{H_l}{W_l}). \tag{5}$$

where $\delta(.)$ is the activation function, H_l represents the output of the l layer, and W is the weight matrix.

The flow of the diagnosis method based on K-means and a graphical neural network is shown in Figure 2. It mainly includes two processes: model training and fault diagnosis. First of all, we need to collect and obtain the data of each state of the circuit, divide it into training sample data and validation

sample data proportionally, and then use the K-means clustering-based pattern layer neuron selection algorithm to filter the training sample data and obtain the final pattern layer neuron sample, and build the graph neural network diagnosis model based on the selected training sample to achieve fault diagnosis. The input of the model is the time-frequency diagram generated from the original vibration signal by wavelet transform and compressed, the size of the diagram is 50 × 50, and the time-frequency diagram contains both time and frequency information, which can help the model to extract the characteristics of the fault signal more effectively. The adjacency matrix, which represents the relationship of the graph data, uses the distance between nodes to capture the relevant dependencies. By inferring the distance between two feature nodes, the nearest k points other than itself are selected to construct the adjacency matrix, which is input into the network to help the GCN layer to extract feature information that depends only on the neighborhood of the data itself.

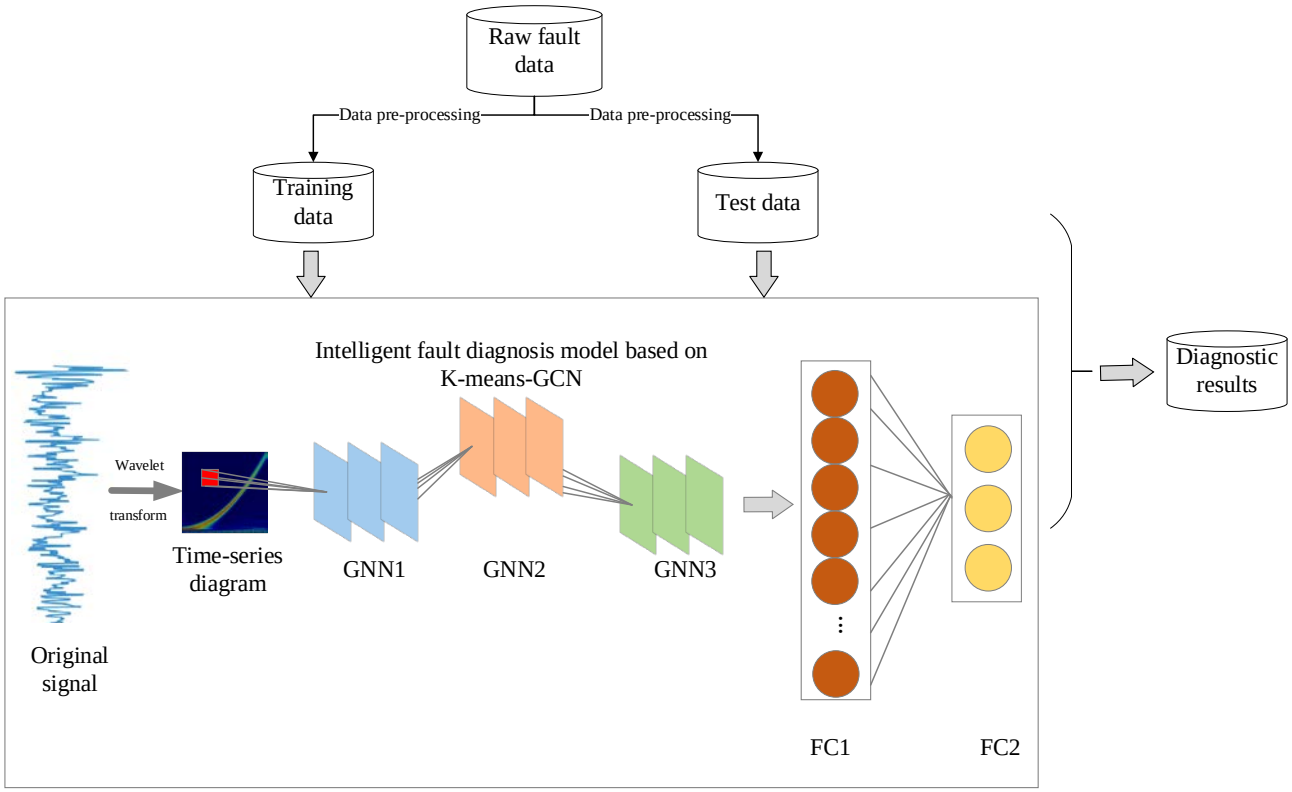

Figure 2. Model of diagnosis method based on K-means clustering and graph neural network.

Dropout is an effective method to prevent overfitting of the network by dropping some random neurons from the network during training to reduce the interdependence between nodes and the co-adaptation relationship between neurons. The model uses Dropout in the fully connected layer and sets Dropout to a common value of 0.5 to randomly discard 50% of neurons, which can reduce the computational overhead of the network and prevent the network from overfitting. The momentum method aims to accelerate learning and solve the problem that the learning parameters in gradient descent are too small or too large, which makes it difficult for the model to converge to the optimal point or the parameters are likely to diverge. The momentum method requires the update of parameters during each state transfer to consider both the current gradient and the historical gradient, making the variable updates in adjacent time steps more consistent in direction and guiding the parameters to converge faster toward the optimal value.

The K-means-based pattern layer neuron preference algorithm mainly performs K-means clustering analysis on the data in the same state, determines the number of clusters K according to the cluster validity index, and finds the cluster centers as the pattern layer neuron samples.

When a fault occurs in the resonant grounding system, the fault current detected at the measurement end of the faulty line has a general tendency to rise within a time window, while the fault current component detected at the measurement end of the non-faulty line has a general tendency to fall within a time window. As shown in Figure 3, it can be seen that there is a greater similarity between the waveforms of the fault line, while there is a greater difference with the sound line.

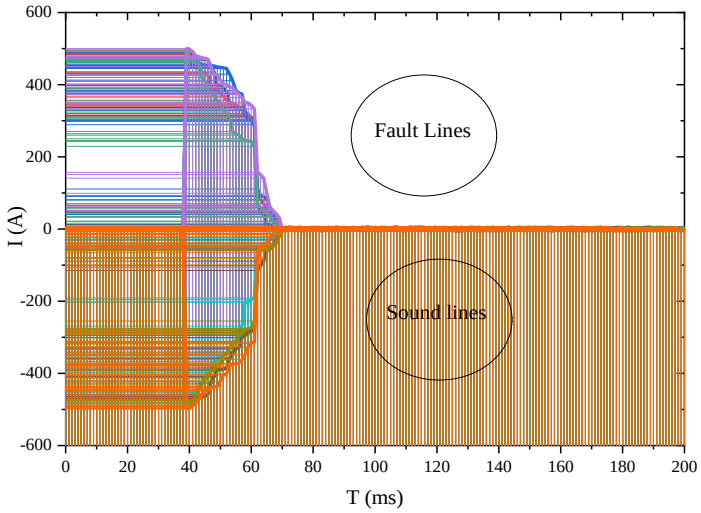

Figure 3. Faulty and non-faulty circuit current curves.

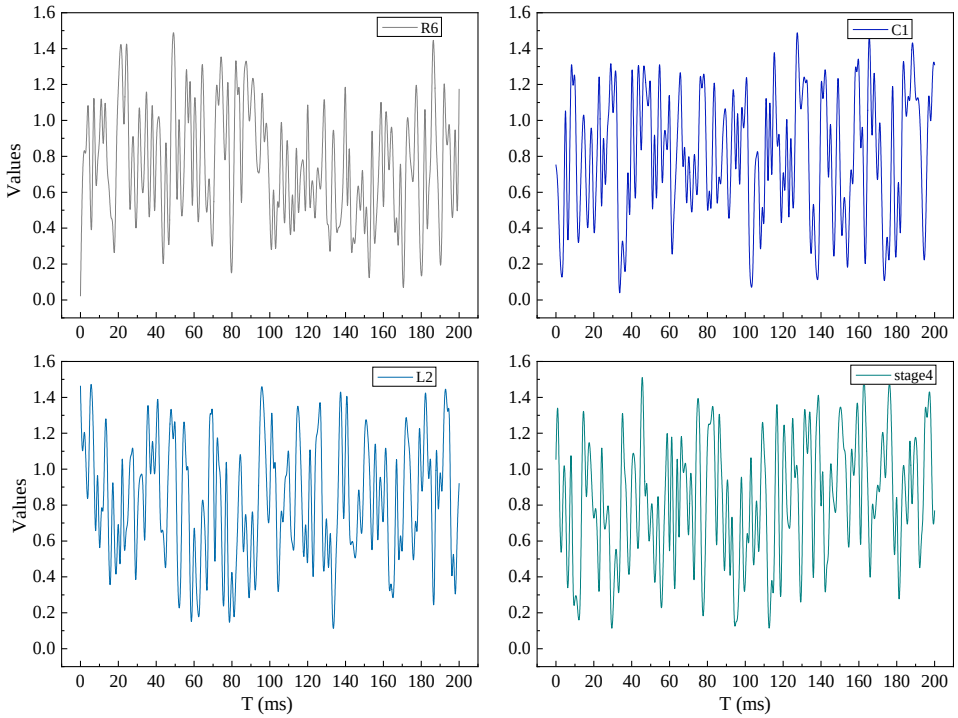

Figure 4. Amplified circuit simulation signal.

In the video amplifier circuit, R6, C1, L2, and the triode were selected as the objects of study, and the voltage signals of the normal circuit; R6, C1, L2, and the triode base intermittent fault; and the triode base permanent fault were collected. In the simulation process, the simulation duration is 4s, the sampling interval is 0.4ms, the number of sampling points is 10000, and 300 Monte Carlo analyses are performed for each fault category, of which 250 groups are training data and 50 groups are test data. The signal acquired at the optimal measurement point of the video amplifier circuit is shown in Fig. 4.

The test samples of the video amplifier circuit are detected according to the detection threshold, and the fault detection curve graph is obtained. From the detection curves, it can be seen that the test samples of each fault category can be classified into the corresponding fault category without false detection. The test data for fault categories 2, 3, 4, and 5 are all intermittent fault states. The effectiveness of the proposed method was verified by conducting experiments on two circuits, which can successfully achieve the determination of intermittent fault signals from the fault signals of analog circuits and can further detect the singularities of intermittent fault signals. The occurrence and recovery of intermittent faults can be determined, thus achieving the accurate analysis of intermittent fault signals of analog circuits.

4. Experimental results and analysis

4.1. Model performance test results

GCNs have become a hot research topic in recent years by their ability to efficiently extract spatial features for learning graph data. Image is a special kind of graph data that can also use GCN to add information from neighboring nodes to the current node to get more complete feature information and exhibit the structure of local connectivity. The biggest difference between image data and other unstructured graph data is that the adjacency matrix can be expressed explicitly in the convolution process. The parameters of the GCN convolution kernel are shared everywhere, which can greatly reduce the number of parameters in each layer of the network and can effectively avoid the phenomenon of overfitting.

In this chapter, the accuracy of fault diagnosis is used to evaluate the proposed GCN semi-supervised learning IC fault diagnosis model, which is trained with the following set of hyperparameters: the number of neighbors K of KNN is 1, the learning rate is 0.001, and the number of iterations is 1000. The model is compared with classical classification algorithms such as the label propagation algorithm (LP, whose kernel function is KNN) and support vector machine (SVM, whose kernel function is RBF). Its kernel function RBF is compared. All methods divide the signal samples into three classes, where the training set trains the signal samples, the validation set selects the parameters for the classification experiments, and the test set tests the classification experiments. As shown in Figure 5, the accuracy of fault diagnosis with different training set ratios is shown in Figure 5.

Compared with LP and SVM, the GCN method proposed in this chapter is the algorithm with the highest accuracy, especially when the size of the labeled samples decreases, the accuracy of the LP algorithm decreases to a larger extent, and the accuracy of the SVM algorithm remains at a lower level. In the case of 10% labeled samples, the proposed method can achieve an accuracy of 95.02%. Therefore, we can conclude that the GCN algorithm can achieve a high recognition accuracy with a

small labeled data size.

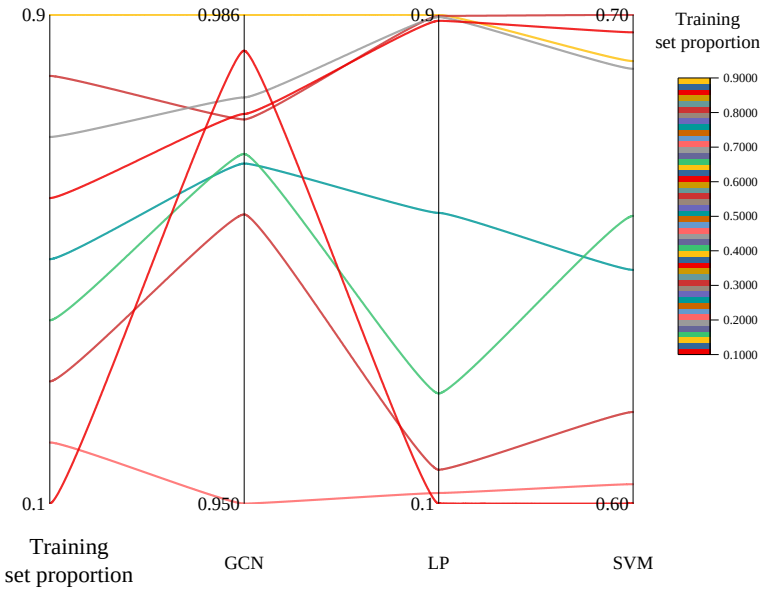

Figure 5. Learning accuracy comparison results.

Using the K-means algorithm, the number of categories of the sample data is determined by continuously traversing the squared error of the distance from the sample points in the classification result to the respective cluster centres for different values of the number of categories in the specified classification range, K-means. When the number of categories is equal to the total number of samples, the squared error is equal to 0. Therefore, it can be concluded that the value of the squared error decreases monotonically with the increase of the K value and approaches 0. When the K value is close to the optimal number of categories, an inflection point of the squared error SSE occurs and the change of the squared error tends to slow down, at which time the K value is the optimal number of categories.

For distribution network fault routing, the sample data should reflect the characteristics of different feeder faults as much as possible. Different fault conditions are set up and the fault locations are traversed every m along the overhead line and every 1 km along the cable line. The sampling rate is 1 Hz. 6-layer wavelet decomposition is performed on the sample data using 10 wavelets to account for the transient zero-sequence current energy at each scale of the wavelet transform within 5 ms after the fault. And calculate the transient zero-sequence current energy under the full frequency band, the energy of line i under the coincident scale j, and as:

$$E_{ij} = \sum_{k=1} D_i^2(k)^2. \tag{6}$$

Whether overhead line fault or electric chicken line fault or cable hybrid line fault, the fault line transient zero sequences current energy and integrated wavelet energy relative bribe are the largest, according to which k-means clustering analysis can be used to transient zero sequences current energy and wavelet relative energy bribe these two dimensions combined to characterize the difference between the fault and not.

The equivariant series scheme, with the number of iterations T determined, uses ε^m as the first term of the series to derive an equivariant series with d as the tolerance and reverses the series as the

privacy budget allocation scheme with ε^m as the privacy budget for the last iteration round, which ensures that the center of mass will not be deformed by the added noise and also allows the earlier iterations or a larger privacy budget.

Assuming that dataset D has 50,000 data and the dimension of the data is 5, the dataset is clustered by DPK-means, the number of k is selected to be 5, the number of iterations T is equal to 10, and the privacy budget is equal to 10. The privacy budget is assigned by the mean and dichotomous methods and the equivariant method, where the minimum privacy budget of the equivariant method is calculated, and the value of p is taken as 0.3 to obtain. The minimum privacy budget value is 0.077, and the three methods are compared as shown in Figure 6.

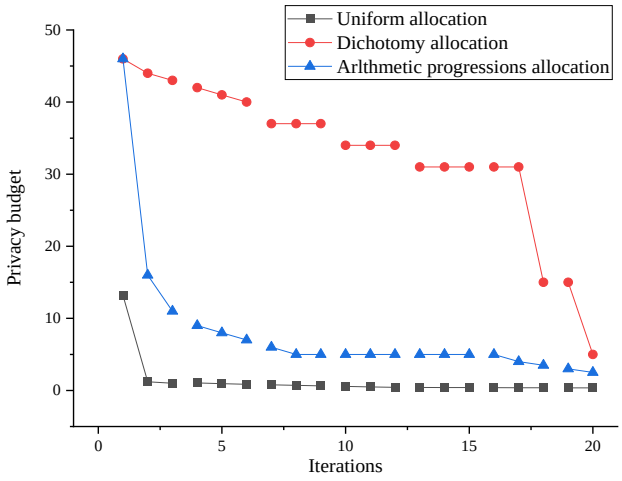

Figure 6. Different privacy budget allocation schemes.

As can be seen from the figure, the mean division method does not consider the influence of the iterative process of DPK-means in different periods on the clustering results. The dichotomous method can allocate more privacy budget in the early stage, but the privacy budget allocated in the later stage is already much smaller than the minimum privacy budget value ε^m, which will have an impact on the clustering effect, and the dichotomous method consumes too much privacy budget in the early stage and cannot exhaust the privacy budget. The dichotomous method consumes too much privacy budget in the early stage and cannot exhaust the privacy budget, which causes a certain degree of waste; the method of equal difference series ensures that the privacy budget assigned each time will not cause the distortion of the centre of mass, but the privacy budget assigned in the early stage is not as good as the dichotomous method. The dimensionality of the data set becomes larger, the upper limit of the number of iterations to be set for K-means clustering will make the tolerance of the difference series smaller and smaller, and the curve of the difference series tends to be flat.

4.2. Digital integrated circuit fault detection simulation experiments

Permanent fault states are obvious and therefore relatively easy to diagnose. Compared with the normal condition, the fault characteristics of intermittent faults are relatively less obvious due to their shorter occurrence time.

It is assumed that the leakage defect test is performed on TSV A and TSV B, where TSV A is

defect-free and TSV B has a leakage defect with a leakage conductance of 20 μSiemens. According to the leakage defect mechanism, the leakage channel causes the voltage on CS to drop faster, resulting in a reduced test response time. Since the test response time required for TSV B is 55 pulses, which is less than the 130 pulses required for a normal TSV A, it can be determined that TSV B has a leakage defect. In practice, the pulse count values for both defective TSVs and defect-free TSVs are randomly shifted due to CMOS process deviations. To verify the test performance of this testable design structure under the process deviation, the following simulation is performed using the Monte Carlo (MC) simulation function of the HSPICE software for the leakage defect test. In the MC simulation, the CMOS threshold voltage and CMOS channel length are made to satisfy the Gaussian distribution of $3\delta_{TH} = 35\%$ and $3\delta_{Leff} = 12\%$. The simulation results are shown in Figure 7, where the horizontal coordinate is 50 Monte Carlo simulations out of 200, and the vertical coordinate is the pulse count result of 20 TSVs during each simulation.

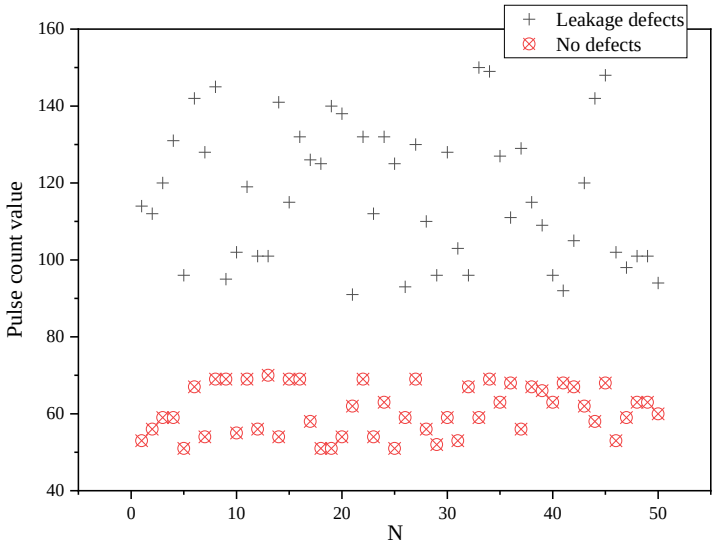

Figure 7. TSV leakage defect test simulation results.

The pulse counts of both the defect-free TSV and the defective TSV with leakage fluctuate to varying degrees in the process deviation case. For each MC simulation, the fluctuations are isotropic and do not confuse the pulse count values of defective and defect-free TSVs. This result is due to the common TSV test cell structure of this measurability design, which allows the process deviation of the test cell to have an isotropic effect on all TSV test results within the group. Therefore, the results of leakage defect detection by the comparative judgment will not be affected by process deviations. The above simulation results show that the method has strong robustness in TSV leakage defect testing.

To verify the overall effectiveness of the digital IC aging fault prediction method with high accuracy, it is necessary to test the digital IC aging fault prediction method with high accuracy. The experimental platform for this test is Matlab, and the digital IC aging fault prediction method with high accuracy, the digital IC aging fault prediction method based on feature information matrix, and the digital IC aging fault prediction method based on matrix perturbation analysis are used respectively. Fault prediction method to predict aging faults in digital integrated circuits, comparing the time used

for prediction by the three different methods, and the test results are shown in Figure 8.

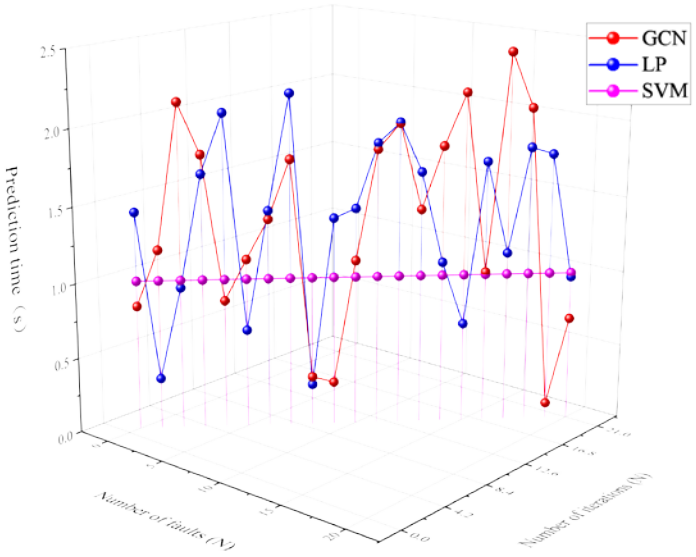

Figure 8. Number of iterations and prediction time graphs.

The time taken to predict the digital IC aging fault in multiple iterations is as high as 24s, which is close to half a minute when the method based on the feature information matrix is used for testing. In the case of the matrix, perturbation analysis-based digital IC aging fault prediction method, the prediction time is within 15s. The digital IC aging fault prediction method uses high-pass filters to remove the noise signals and redundant signals present in the digital IC, which reduces the time used for prediction and verifies the high prediction efficiency of the digital IC aging fault prediction method.

5. Conclusion

In this paper, we propose a fault diagnosis method based on the combination of K-means and a graphical neural network for analog circuits. Through the comprehensive analysis of K-means clustering and clustering validity indexes, we give the pattern layer neuron preference process, screen the suitable training sample data as the pattern layer neurons of a graphical neural network to participate in training, and achieve high fault accuracy fault diagnosis with fewer training samples. This reduces the complexity of the model and saves the fault diagnosis time. In the study of TSV defect modeling, the RLGC parameter equivalent circuit models of TSV void defect, open circuit defect, leakage defect, micro liner unaligned defect, and micro liner missing defect are established. With this equivalent circuit model, the mathematical mapping between the physical parameters of each TSV defect and the electrical parameters of the equivalent model is established. By comparing and analyzing the active filter circuit with the traditional graph neural network method and the random graph neural network method, the superior performance of the method in terms of fault diagnosis accuracy and fault diagnosis efficiency is verified. In this paper, certain achievements are made in fault monitoring of digital integrated circuits, but the accuracy of model parameter training can be improved. The fault diagnosis method can be extended to different application scenarios, and for specific application scenarios, other specific algorithms with more specific targeting can be considered, or multiple optimization algorithms can be integrated to obtain the optimal training strategy.

Acknowledgments

This work is supported by Shandong University.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. Z. Guo, Y. Shen, S. Wan, W. Shang, K. Yu, Hybrid intelligence-driven medical image recognition for remote patient diagnosis in internet of medical things, *IEEE J. Biomed. Health Inform.*, **26** (2022), 5817–5828. https://doi.org/10.1109/JBHI.2021.3139541
2. Q. Zhang, Z. Guo, Y. Zhu, P. Vijayakumar, A. Castiglione, B. B. Gupta, A deep learning-based fast fake news detection model for cyber-physical social services, *Pattern Recogn. Letters*, **168** (2023), 31–38. https://doi.org/10.1016/j.patrec.2023.02.026
3. X. Shen, G. Shi, H. Ren, W. Zhang, Biomimetic vision for zoom object detection based on improved vertical grid number YOLO algorithm, *Front. Bioeng. Biotechnol.*, **847** (2022), 905583. https://doi.org/10.3389/fbioe.2022.905583
4. Z. Guo, D. Meng, C. Chakraborty, X. Fan, A. Bhardwaj, K. Yu, Autonomous behavioral decision for vehicular agents based on cyber-physical social intelligence, *IEEE Transact. Comput. Soc. Syst.*, (2022), 1–12. https://doi.org/10.1109/TCSS.2022.3212864
5. S. R. Gaddam, V. V. Phoha, K. S. Balagani, K-means$^+$ ID3: A novel method for supervised anomaly detection by cascading K-means clustering and ID3 decision tree learning methods, *IEEE Transact. Knowl. Data Eng.*, **19** (2007), 345–354. https://doi.org/10.1109/tkde.2007.44
6. Q. Li, L. Liu, Z. Guo, P. Vijayakumar, F. Taghizadeh-Hesary, et al., Smart assessment and forecasting framework for healthy development index in urban cities, *Cities*, **131** (2022), 103971. https://doi.org/10.1016/j.cities.2022.103971
7. H. Gong, S. Cheng, Z. Chen, Q. Li, C. Quilodrán-Casas, D. Xiao, et al., Efficient digital twin based on machine learning SVD autoencoder and generalised latent assimilation for nuclear reactor physics, *Ann. Nuclear Energy*, **179** (2022), 109431. https://doi.org/10.1016/j.anucene.2022.109431
8. Z. Guo, K. Yu, A. Jolfaei, G. Li, F. Ding, A. Beheshti, Mixed graph neural network-based fake news detection for sustainable vehicular social networks, *IEEE Transact. Intell. Transport. Syst.*, (2022). https://doi.org/10.1109/TITS.2022.3185013
9. Y. Zuo, W. Su, S. Zhang, S. Wang, C. Wu, L. Yang, et al., Discrimination of membrane transporter protein types using K-nearest neighbor method derived from the similarity distance of total diversity measure, *Molecul. bioSyst.*, **11** (2015), 950–957. https://doi.org/10.1039/c4mb00681j
10. X. Wang, X. Zhu, M. Ye, Y. Wang, C. D. Li, Y. Xiong, et al., STS-NLSP: A network-based label space partition method for predicting the specificity of membrane transporter substrates using a hybrid feature of structural and semantic similarity, *Front. Bioeng. Biotechnol.*, **7** (2019), 306. https://doi.org/10.3389/fbioe.2019.00306
11. R. Isermann, Supervision, fault-detection and fault-diagnosis methods—an introduction, *Control Eng. Pract.*, **5** (1997), 639–652. https://doi.org/10.1016/S0967-0661(97)00046-4

12. J. Wang, A. Ma, Y. Chang, J. Gong, Y. Jiang, R. Qi, et al., scGNN is a novel graph neural network framework for single-cell RNA-Seq analyses, *Nat. Commun.*, **12** (2021), 1–11. https://doi.org/10.1038/s41467-021-22197-x
13. H. Hong, H. Guo, Y. Lin, X. Yang, Z. Li, J. Ye, An attention-based graph neural network for heterogeneous structural learning, in *Proceedings of the AAAI conference on artificial intelligence*, **34 (**2020), 4132–4139.
14. O. Mohammadrezapour, O. Kisi, F. Pourahmad, Fuzzy c-means and K-means clustering with genetic algorithm for identification of homogeneous regions of groundwater quality, *Neural Comput. Appl.*, **32** (2020), 3763–3775. https://doi.org/10.1007/s00521-018-3768-7
15. J. Wang, J. Yang, D. Chen, L. Jin, Y. Li, Y. Zhang, Gas detection microsystem with MEMS gas sensor and integrated circuit, *IEEE Sensors J.*, **18** (2018), 6765–6773. https://doi.org/10.1109/JSEN.2018.2829742
16. D. Susanto, Komparator Integrated Circuit Digital, *Jurnal Instrumentasi dan Teknologi Kebumian*, **1** (2022), 32–39.
17. P. Ghosh, R. S Chakraborty, Recycled and remarked counterfeit integrated circuit detection by image-processing-based package texture and indent analysis, *IEEE Transact. Industr. Inform.*, **15** (2018), 1966–1974. https://doi.org/10.1109/TII.2018.2860953
18. S. Herbst, G. Rutsch, W. Ecker, M. Horowitz, An open-source framework for FPGA emulation of analog/mixed-signal integrated circuit designs, *IEEE Transact. Computer-Aided Design Integr. Circ. Syst.*, **41** (2021), 2223–2236. https://doi.org/10.1109/tcad.2021.3102516
19. H. Gong, S. Cheng, Z. Chen, Q. Li, Data-enabled physics-informed machine learning for reduced-order modeling digital twin: Application to nuclear reactor physics, *Nuclear Sci. Eng.*, **196** (2022), 668–693. https://doi.org/10.1080/00295639.2021.2014752
20. L. Feng, C. Zhao, Fault description-based attribute transfer for zero-sample industrial fault diagnosis, *IEEE Transact.Industr. Inform.*, **17** (2020), 1852–1862. https://doi.org/10.1109/TII.2020.2988208
21. A. Mellit, S. Kalogirou, Handbook of Artificial Intelligence Techniques in Photovoltaic Systems: Modeling, Control, Optimization, Forecasting and Fault Diagnosis, Academic Press, Elsevier, 2022.
22. F. Xiao, Z. Cao, A. Jolfaei, A novel conflict measurement in decision-making and its application in fault diagnosis, *IEEE Transact. Fuzzy Syst.*, **29** (2020), 186–197. https://doi.org/10.1109/TFUZZ.2020.3002431
23. Y. Lei, Z. He, Y. Zi, Application of an intelligent classification method to mechanical fault diagnosis, *Expert Syst. Appl.*, **36** (2009), 9941-9948. https://doi.org/10.1016/j.eswa.2009.01.065
24. Z. Guo, K. Yu, A. K. Bashir, D. Zhang, Y. D. Al-Otaibi, M. Guizani, Deep information fusion-driven POI scheduling for mobile social networks, *IEEE Network*, **36** (2022), 210-216. https://doi.org/10.1109/MNET.102.2100394