A Study on Reducing Transistor Electrical Characteristic Inspection Processing Time Using Machine Learning

Hyungjin Lee, Daehong Kim, Jungmoon Lee

Display Research Center, Samsung Display Co., Ltd., Giheung-gu, Yongin-City, Gyeonggi-Do, Korea

Abstract

With growing competition in display manufacturing, ensuring the quality of TFT components is crucial. This study uses machine learning to reduce the Tact Time of TFT electrical inspections while maintaining accuracy. By analyzing data from 47,160 glass substrates, clustering and predictive modeling techniques were applied. These methods improved inspection efficiency, enhanced production yield, and reduced costs, providing a scalable solution for modern manufacturing.

Author Keywords

Transistor Inspection; Machine Learning; Processing Time Reduction; I-V data analysis

1. Introduction

The recent rapid advancements in display technology have created a highly competitive environment, standardizing process benchmarks among suppliers [1]. Display manufacturers are adopting a variety of advanced process technologies to enhance display driving characteristics [2]. Transistor components are critical to display performance, and detecting degradation through electrical characteristic testing is crucial for improving overall yield and enhancing competitiveness [3].

Advancements in display technology demand high functionality, which has led to an increase in process complexity. These changes require new approaches to inspection processes. Balancing process speed and quality has become a key challenge for manufacturers, and machine learning is being highlighted as a promising tool to address this issue.

The electrical inspection of transistors typically involves measuring the current between the drain and source while increasing the gate voltage in 0.25V increments [4]. However, such traditional methods require considerable time to conduct thorough testing of all characteristics. Especially for tests aimed at detecting image retention issues, repeated testing in both forward and reverse directions effectively doubles the inspection time. Thus, reducing the processing time for electrical characteristic testing has become a pressing issue in the industry. Thus, reducing the processing time for electrical characteristic inspection has become a pressing issue in the industry.

As display resolutions increase and component complexity grows, the amount of data handled during each inspection is also expanding [5]. These temporal and resource limitations render traditional inspection methods increasingly inefficient [6]. This research proposes a new solution to enhance inspection efficiency by using machine learning, thereby boosting inspection capacity while ensuring consistent product quality. Machine learning offers a powerful tool for analyzing large datasets and identifying critical patterns, enabling more focused and expedited inspection processes [7]. By adopting a data-driven approach, this study aims to address the need for faster inspection methods in modern display manufacturing and simultaneously improve the efficiency and quality of manufacturing processes.

2. Materials and Method

2.1 Data Collection and Preprocessing

This study utilized I-V data obtained from 47,160 units of 2nd generation glass substrates collected in a production environment. This dataset included current measurements from each transistor. Prior to applying machine learning algorithms, preprocessing steps including outlier removal and handling of missing values were conducted to ensure data quality and reliability [8]. Preprocessing was critical to optimizing the machine learning model's performance, and the key steps were as follows:

Outlier Detection and Removal: Due to the high sensitivity of the characteristic inspection probe pin, environmental factors can generate numerous outliers [9]. These significantly impact data quality and can negatively affect model training and prediction accuracy. Therefore, outlier detection and removal were essential steps. We used statistical techniques such as Interquartile Range (IQR) analysis to identify outliers and minimize their impact on the model. Values beyond Q1 - 1.5 * IQR or Q3 + 1.5 * IQR were identified as outliers and removed [10]. This analysis reduced data distortion, leading to more accurate and reliable results.

Data Normalization: Measurement values are usually on a logarithmic scale with large variations between individual values. These variations can degrade the performance of learning algorithms, necessitating a normalization process. By standardizing the input range of the data, normalization improved model convergence speed and stabilized the learning process. Particularly, normalization played a crucial role in maintaining consistency between features when there were significant differences in unit scales. This prevented the model from biasing towards specific features and ensured a uniform reflection of the overall data distribution [11]. By performing normalization, we enhanced the model's generalization performance and reduced prediction errors, thereby contributing to effective learning outcomes across different machine learning models.

2.2 Relative Importance Evaluation of Measurement Nodes and Key Node Selection

The relative importance of each measurement node was assessed by analyzing the variation and differential values among nodes or by using various machine learning techniques to estimate the importance of each node [12]. This allowed us to quantify the relative importance of each node's contribution to the I-V curve and select the most critical nodes.

Density-Based Evaluation Methods: Density-based methods, including K-means, Density-Based Spatial Clustering of Applications with Noise (DBScan), and Affinity Propagation (AP), were employed to evaluate node importance considering data distribution and proximity among nodes. These methods interpreted the data distribution differently, providing diverse perspectives on node importance [13] [14].

ISSN 0097-996X/25/5601-0644-$1.00 © 2025

Graph Neural Network Evaluation Methods: Graph Neural Network (GNN) approaches, including Graph Convolutional Network (GCN), Graph Attention Network (GAT), and Graph Sample and Aggregation (GraphSAGE), were also utilized. These techniques modeled complex interactions among nodes based on graph structure, providing a refined understanding of inter-node relationships [15] [16].

Evaluation Metric (Cosine Similarity): The results from each method were compared using cosine similarity, which measures directional similarity between two vectors by the angle between them. Given the high correlation among adjacent measurement data, cosine similarity proved useful in evaluating the overall directionality of the graph [17].

2.3 Predictive Modeling for Unmeasured Nodes

To predict the current values at unmeasured nodes based on the selected key nodes, various regression models were applied and compared. Extreme Gradient Boosting (XGBoost), for instance, enhances prediction performance through iterative boosting, effectively learning complex patterns. Support Vector Machine (SVM) is well-suited for classification and regression of high-dimensional data, utilizing appropriate kernels to solve nonlinear problems. Linear regression is simple and effective for modeling linear relationships, whereas Ridge regression uses L2 regularization to mitigate overfitting. Gradient Bossting Machine (GBM) gradually reduces errors through repetitive learning, thereby improving model performance. Random Forest, on the other hand, combines multiple decision trees to enhance prediction stability and accuracy. Polynomial interpolation was also used to smoothly estimate values between known data points, making it advantageous for capturing nonlinear trends with limited data [18] [19].

3. Results and Discussion

3.1 Selection of Key Measurement Nodes

Analyzing the cosine similarity results revealed significant differences in the performance of density-based evaluation methods and GNN methods. As shown in Table 1, K-means achieved the highest cosine similarity value of 0.992, indicating its strength in reflecting cluster characteristics and effectively selecting key nodes. On the other hand, DBScan also showed high similarities of 0.972, respectively, suggesting that approaches based on data proximity or density effectively retain critical information from the I-V curves. AP demonstrated a similarity of 0.978, confirming its effectiveness in capturing representative features among the data.

Among GNN methods, GraphSAGE achieved the best performance with a similarity of 0.944, while GCN and GAT recorded lower similarities of 0.895 and 0.885, respectively. These results suggest that graph-based approaches require more information and interactions between nodes to effectively model complex relationships. Particularly, GAT showed relatively low performance, which could indicate that the attention mechanism might not have sufficiently captured the structural characteristics needed for effective node importance learning.

Overall, density-based evaluation methods outperformed GNN techniques, indicating that the given I-V data's characteristics are effectively analyzed through clustering structures and proximity-based approaches. As shown in Figure 1, GNN techniques have demonstrated limitations in identifying key nodes in tightly connected inflection area of data. These results provide insights into the efficient reduction of node numbers while retaining critical information, suggesting that density-based approaches are preferable when formulating future node selection strategies.

Table 1. Comparison of Models for Node Selection Using Cosine Similarity Metric

	Kmeans	DBScan	AP	Graph SAGH	GCN	GAT
C/S	0.992	0.972	0.978	0.944	0.895	0.885

3.2 Predicting Unmeasured Nodes

The predictive performance of the models used to estimate the current values at unmeasured nodes was compared based on Mean Squared Error (MSE). The lowest MSE was achieved by the quadratic interpolation method, recording an MSE of 0.00044, indicating that this method is most suitable for smoothly connecting unmeasured segments and capturing nonlinear trends. Cubic interpolation also performed well, with an MSE of 0.00074, although slightly higher than the quadratic method. Among regression models, Gradient Boosting Machine (GBM) showed the best performance with an MSE of 0.0024.

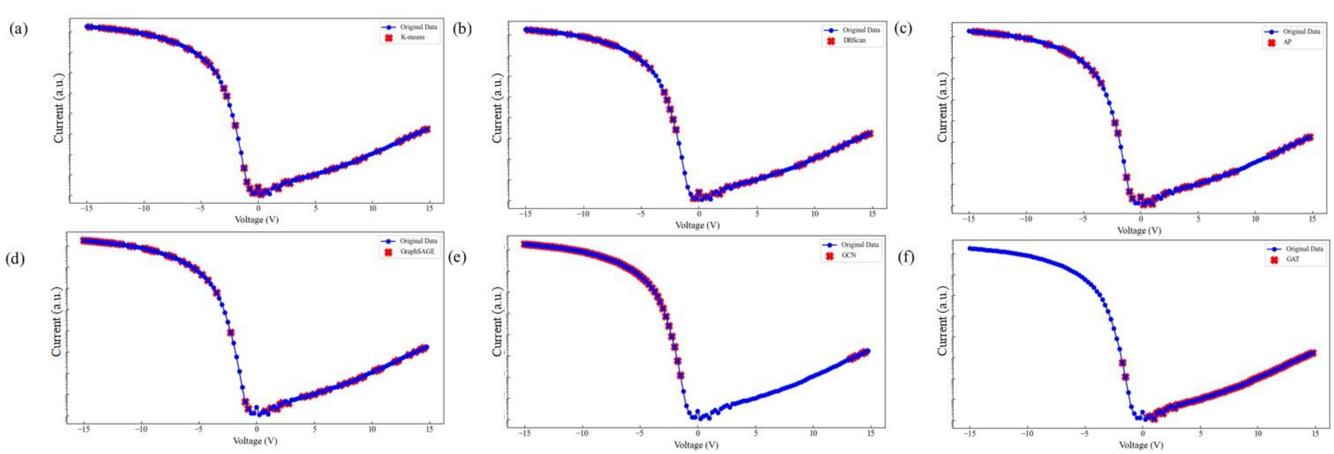

Figure 1. Selection of Key Nodes using (a) K-means (b) DBScan (c) AP (d) GraphSAGE (e) GCN (f) GAT

Table 2. Comparison of Prediction Model Using MSE

	XGBoost	SVM	Ridge Regression	GBM	Random Forest	1D Interpolation	2D Interpolation	3D Interpolation
MSE	0.024	0.014	0.142	0.0024	0.0035	0.112	0.00044	0.00074

As shown in Table 2, SVM and XGBoost recorded MSEs of 0.014 and 0.024, respectively, reflecting their strengths in effectively learning complex patterns in high-dimensional data. In contrast, Ridge regression showed a relatively high error of 0.142, indicating that simple linear models were insufficient to capture the nonlinear characteristics of the data. Linear interpolation also recorded an MSE of 0.112, performing worse compared to polynomial methods.

These results confirm that polynomial interpolation, particularly the quadratic method, provides the best performance for predicting unmeasured nodes. Furthermore, boosting-based methods like GBM also demonstrated promising predictive power, making them viable alternatives. Ultimately, model selection should consider not only predictive accuracy but also computational complexity and efficiency in real-world environments.

3.3 Increased Monitoring Capacity thorough Processing Time Reduction

The system was implemented as shown in Fig. 2. The series of processes such as transistor inspection, key node selection, and the prediction of unmeasured nodes were automated and applied to the manufacturing system.

Reducing the processing time of inspections through machine learning significantly increased the number of samples that could be monitored on the production line. This allowed for the early detection of degradation and image retention issues, contributing to quality improvements. The proposed method simplified the inspection process while maintaining accuracy, thus enhancing overall quality monitoring efficiency. Additionally, automation of the inspection reduced dependency on human resources and reinforced consistency in inspection outcomes.

Reducing processing time is particularly advantageous in mass production environments, as it can lead to increased overall production volumes. Such reductions also contribute to cost savings by enabling the early identification and isolation of defective products. Moreover, the capability to predict values at unmeasured nodes allows manufacturers to identify potential issues early in the production cycle, preventing defective products from proceeding to subsequent stages.

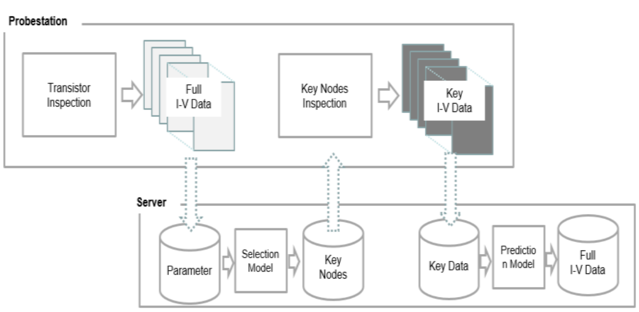

Figure 2 Workflow of the Implemented ML Model

4. Conclusion

This study proposed a methodology to reduce the processing time of electrical characteristic inspections for the transistor by applying machine learning techniques. By analyzing I-V data collected from 47,160 glass substrates and employing various modeling approaches, we identified key measurement nodes to enhance inspection efficiency. The results demonstrated that it is possible to reduce the number of measurement nodes while maintaining a high level of accuracy in electrical characteristic inspections. Using a linear regression model, we successfully predicted the current values at unmeasured nodes, ultimately achieving a high degree of similarity with the original I-V curves.

This methodology significantly improved inspection efficiency without compromising inspection quality. As a result, manufacturers can reduce the processing time, expand monitoring capacity, and improve overall production quality through early defect detection.

Future research will focus on evaluating additional machine learning models to further enhance predictive accuracy and integrating these methodologies into real-time monitoring systems in manufacturing processes. By advancing the automation of transistor inspection processes, we aim to contribute to overall efficiency and quality improvement in the display industry.

5. References

1. Lee, J., & Kim, S. Advanced Display Technologies and Industry Trends. *Journal of Display Science*, 2022, 15(4): 345-367.

2. Smith, R. J. Advanced Process Technologies for Display Performance Enhancement. *Journal of Applied Electronics*, 2021, 32(7): 501-515

3. Liu, M., & Zhang, W. Importance of TFT Inspection in Early Degradation Detection. *Display Electronics Journal*, 2020, 18(6): 145-162.

4. Evans, C. Electrical Characteristic Inspections for TFT Devices. *Journal of Electronic Testing*, 2022, 40(1): 78-91.

5. Wang, Y., & Li, X. Data Complexity in High-Resolution Display Inspections. *Electronics and Automation*, 2022, 45(3): 315-328.

6. Johnson, T. A. Optimizing Inspection Processes in TFT Manufacturing. *International Journal of Manufacturing Systems*, 2020, 29(2): 225-238.

7. Park, H., & Choi, K. Machine Learning Applications in Display Manufacturing. *IEEE Transactions on Electronics*, 2021, 58(3): 1125-1138.

8. Zhang, L. Data Preprocessing Techniques for Reliable Machine Learning. *Journal of Data Analytics*, 2021, 11(3): 211-223.

9. Leshem, S., Cohen, N., Pham, S., & Lim, M. TFT Array: Inspection, Testing, and Repair. *Flat Panel Display Manufacturing Technology*, 2018, Wiley Online Library.

10. Sesia, I., Cantoni, E., & Cernigliaro, A. An efficient and configurable preprocessing algorithm to improve stability analysis. *IEEE Transactions on Instrumentation and Measurement*, 2015, *64*(7): 1921–1930.

11. Singh, D., & Singh, B. Feature-wise Normalization: An Effective Way of Normalizing Data. *Pattern Recognition*, 2022, 122: 108266.

12. Zhang, M., Wang, X., Jin, L., Song, M., & Li, Z. A new approach for evaluating node importance in complex networks via deep learning methods. *Neurocomputing*, 2022, 491: 320–331.

13. Chen, Y., Zhou, L., Pei, S., Yu, Z., & Chen, Y. KNN-BLOCK DBSCAN: Fast clustering for large-scale data. *IEEE Transactions on Knowledge and Data Engineering*, 2019.

14. Fuchs, M., & Höpken, W. Clustering: Hierarchical, k-Means, DBSCAN. *In: Innovative Approaches, Methodologies, and Applications in Knowledge Discovery and Information Analysis*, Springer, 2022.

15. Yang, H. GatedGCN with GraphSage to Solve Traveling Salesman Problem. *International Conference on Artificial Neural Networks*, Springer, 2023.

16. Chang, L., & Branco, P. Graph-based solutions with residuals for intrusion detection: The modified e-graphsage and e-resgat algorithms. *arXiv preprint arXiv:2111.13597*, 2021.

17. Minmin, W., Shengli, S., & Yejin, L. A compressive tracking method based on Gaussian differential graph and weighted cosine similarity metric. *IEEE Signal Processing Letters*, 2016, *23*(11): 1612–1616.

18. Ferreira, A. J. Boosting the Prediction of Extreme Values. *Repositório Aberto da Universidade do Porto*, 2022.

19. Nemer, Z. N. Oil and Gas Production Forecasting Using Decision Trees, Random Forest, and XGBoost. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 2024