TEG Electrical Virtual Measurement and Monitoring Based on Interpretable Machine Learning Method

Peng Li, Jing Ba*, Zhiwei Tan*, Yangxing Liu****
***TCL China Star Optoelectronics Technology Co., Ltd., Shenzhen, China**
****Wuhan TCL Research Co., Ltd., Wuhan, China**

Abstract
In the process of semiconductor display manufacturing, many testing devices are used to perform sampling inspection of panel characteristics, monitor the quality of the production process, and ensure the product yield. With the development of artificial intelligence (AI), data-driven virtual Measurement (VM) AI models can predict the characteristics of display backplane devices in real time based on process parameters, and predict the device characteristics of display backplane devices in advance (usually > 2 days). In this paper, an AI prediction method based on the fusion of random forest and deep learning is designed, which can be used to predict the on-state current Ion & threshold voltage Vth values that affect the performance of the display panel in real time. The prediction accuracy of our proposed model is more than 95%, the MAPE is less than 10%, and the R^2 is between 0.5 and 0.6. Also, the proposed model has good interpretability, and can assist in locating the candidate parameter factors that cause TEG values to exceed control. The model is deployed in the production environment and linked to the SPC system, which can achieve 100% virtual measurement and anomaly monitoring and interception. The model can be automatically updated regularly on the platform to alleviate the decline of prediction effect caused by equipment data drift and model degradation.

Author Keywords
Virtual measurement, Deep learning, Ensemble tree model, Interpretability, Automatic update.

1. Introduction
In the manufacturing process of semiconductor display, many testing devices are needed for sampling inspection of panel characteristics to realize quality control in the production process and ensure product yield[1]. Most Advanced Process Control (APC) tools strongly rely on physical measurement tools. For example, the electrical properties of TFT devices are measured by regular/quantitative random sampling after each processing step, which means that the quality of the production panel is unknown between two measurements. When equipment failure or anomaly is not warned in time, many defective panels may be produced between measurements, resulting in a large number of scrapped panels, which seriously affects production costs and yield. At present, in the production process of display panels, the array section process is complex, and there are many production testing sites. TEG (Test Element Group) electrical testing is an important monitoring step that affects the performance of the panel, but limited by the production laoding, the TGE electrical testing rate of the production line is about 0.3%. Among them, the TEG test items include IV characteristics, namely current-voltage characteristic curve test and capacitance and resistance measurement. If the IV electrical characteristics measured by TEG drift or exceed the gauge, the product will show abnormal display and cannot be repaired.

An effective method to solve the above problems is to use the process parameter data of production equipment to predict the characteristic value of each panel, and realize the virtual measurement of the characteristic value. The goal of virtual measurement is to develop a robust prediction that can provide measurement estimates and be able to handle process drift and step function changes caused by preventive maintenance disturbances. Virtual measurement technology can be used to monitor the production process data and predict the product quality or process, replacing the traditional offline and delayed quality sampling inspection with online and immediate quality inspection. This method uses the data of the FDC system, which collects the data of sensor devices for each process run online. These data are called process variables or FDC data[2]. The AI virtual measurement prediction algorithm model is established based on the production equipment parameter data of the FDC system and the measured values of the TEG test equipment, and the on-state current Ion & threshold voltage Vth are 100% predicted and monitored to reduce the risk of product anomalies.

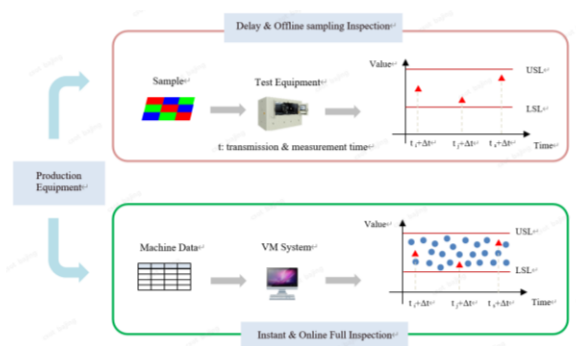


Figure 1. Schematic diagram of actural measurement and Vmmonitoring of semiconductor display panel characteristics.

2. Methodologies

(a) Data acquisition
Based on the analysis of TEG electrical related production processes CVD/DRY/WET, six key sites were locked, and the time series data source of FDC system was obtained. The parameter dimension of six sites was > 2K, and semi-annual panel Glass was obtained for modeling training.

Table 1. Description of FDC process site parameters.

Process	Site	Describe
CVD	NO.1	Insulation layer GI and Amorphous Silicon and N+ Depo
	NO.2	Insulation layer PV1 Depo
DRY	NO.3	1st Dry Etch Affects the thickness of GI remain

ISSN 0097-996X/25/5602-1036-$1.00 © 2025

	NO.4	2nd Dry Etch As the tail of AS and N+ decrease, the effective length of channel increase, while both Ion&Ioff decrease
WET	NO.5	1st Wet Etch Adjust the shape of silicon island
	NO.6	2nd Wet Etch Affects the length of N+ tail and channel length

(b) Data preparation and feature engineering

1. The original FDC data is the time series data of three processes: CVD, DRY and WET. The specific process parameters of each process pass through different sites, and the sites have different process steps (stepname). In this method, the characteristics of each parameter are extracted according to the process data of CVD, DRY and WET processes. The target value is the Vth or Ion value of 16 measurement points on a panel.

2. Because of the near thousands of features' dimension, we used feature filtering based on statistical information and feature filtering based on model. Specifically, we used the correlation between features and target values (spearman and pearson) for feature filtering, and retained the features with high correlation between target values and features. The Maximum Mutual Information Coffiencent (MIC) between the feature and the target value was calculated, and a certain proportion of features with high MIC coefficients were retained.

3. Model based adversarial verification and feature importance is used for feature screening Specifically, the data was randomly divided into training set and test set for binary classification, and then the classification modeling was carried out based on each feature at a time, and the features that made the classification accuracy greater than a certain proportion were deleted. For the retained features, the ensemble tree model was used to model the regression prediction model, and the feature importance coefficient was analyzed for the modeled features, and a certain proportion of important features were retained.

4. In the real source data, the proportion of over-specification target values is low. If it is directly used to train the model, the prediction accuracy of the over-specification part of the model will be low. In order to alleviate this problem, this paper uses SMOTE[3] method to enhance the out-of-gauge samples other than 3sigma, and divides the samples based on SPXY method in the training process.

(c) Model design and training

The proposed method uses ensemble learning random Mori regression[4] and residual deep network[5]. This method needs to predict the value of 16 points, so the random forest model is designed to use the predicted point ID as a category feature and added to the model training process. The model can predict different point values at one time, which improves the real-time performance of model prediction. At the same time, the Bayesian optimization method is used to optimize the random forest hyperparameters in the training process of the random forest, so that the obtained model is optimal. In order to enhance the robustness of prediction, and due to the inherent characteristics of the ensemble tree model, the predicted results are closer to the central value, this paper designs a deep residual network model to predict the electrical values of 16 points, that is, the final fully connected layer directly outputs the values of 16 points, which makes the model have better prediction effect at the 3sigma boundary and stronger generalization ability.

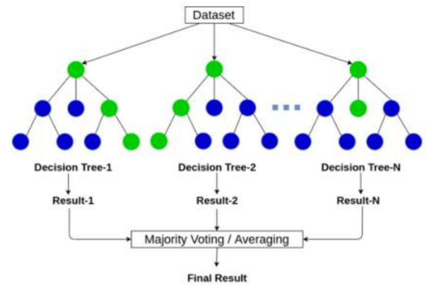


Figure 2. Random Forest Model.

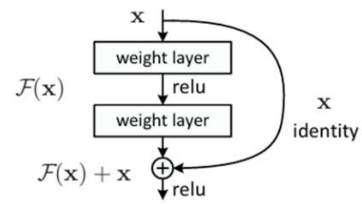


Figure 3. Residual learning: a building block.

$$Loss(y, y') = \frac{1}{N}\sum_i^N \mathbb{I}_{y_i' \geq y_i}(1-\gamma)|y_i - y_i'| + \mathbb{I}_{y_i' < y_i}\gamma|y_i - y_i'| \quad (1)$$

where：y_i is the actural value， y_i' is the predicted value， $\gamma\epsilon(0,1)$ is the quantile level.

Five-fold cross validation was used in the training process of the random forest model, and the model with the best cross validation effect was selected as the final prediction model. At the same time, the quantile loss function (Formula 1) is used in the deep residual network, the Adam gradient descent method is used to train the model, and the Cosine annealing attenuation strategy is used.

(d) Self-update of the model

With the progress of production, the process parameters of the equipment will drift, and the sampling inspection of the product has been carried out continuously. The process parameters drift will lead to the decline of the model prediction effect. The method proposed in this paper will regularly use the accumulated data to update the model. Specifically, when the accuracy index of the online model prediction results on the sampling data is lower than the set threshold index for a period of time, the model will start automatic update. When the accuracy index of the re-trained model is higher than that of the old model, the old model will be updated to complete the self-update of the model. The update process is shown in Figure 7.

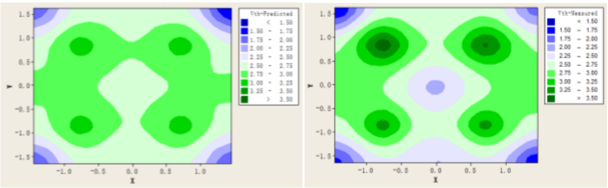
(e) Interpretability

The virtual measurement method designed in this paper not only has advantages in prediction effect, but also has certain interpretability. When the batch abnormality of electrical Vth or Ion occurs, the tree model can assist in the analysis of adverse factors based on Shap[6] interpretability values. The deep learning model can use the Captum[7] framework and Shap input feature factor feature importance, directional gradient and sensitivity analysis, from the perspective of data to assist business personnel to analyze and improve the adverse factors.

3. Predict Results Analysis

3.1. Single slice results

Based on the separate prediction of 16 points per large plate, each large plate of G11 generation line can be cut into 8 pieces of 65-inch display screens, and the TSK of TEG is distributed around each panel. The prediction and measured data match well. Figure 4 shows the contour map of Vth prediction and measured data comparison.



(a) Actual measured and (b) Predicted results of Vth

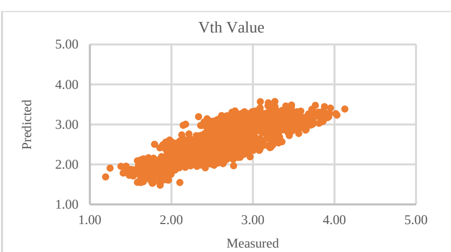
Figure 4. Comparsion of actural &predicted data of different points.

3.2. Batch prediction results and experiments

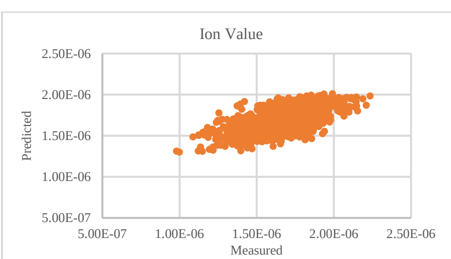
Based on the production data of 80+ glass and thousands of points, the blind test is carried out, and the data is measured and predicted. The prediction results MAPE, R^2 and other indicators are used to evaluate the overall performance of the machine learning model. MAPE stands for Absolute percentage error. It is a performance measure used to evaluate the level of difference between predicted and measured values. The smaller the value, the better. R^2 is the coefficient of determination, an indicator of how well the model fits, and it ranges from 0 to 1, with closer to 1 being better. MAPE and R^2 are given in Equation (2) and Equation (3), respectively.

$$MAPE = \frac{1}{n} \times \sum_{i=1}^{n} abs\left(\frac{y_i - y_i^{'}}{y_i}\right) \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_i^{'})^2}{\sum_{n}^{i=1}(y_i - \bar{y})^2} \quad (3)$$



(a) Predicted results and measured results of Vth



(b) Predicted results and measured results of Ion

Figure 5. Comparsion of actural and predicted data of different panels.

As Table 2 and Table 3 depicted, the fusion model in this paper predicts MAPE < 10% and $R^2 > 0.5$. Based on the prediction ability of the model, the single point accuracy criterion was set, and the difference between Ion prediction and measurement was less than 15%, and the difference between Vth prediction and measurement was less than 0.6V, which was defined as the acceptable prediction accuracy. Based on these two items, the prediction accuracy of the model was > 95%.

Table 2. Comparsion of different models' MAPE and R^2.

Electrical properties	Ion		Vth	
	MAPE	R^2	MAPE	R^2
Tree-Based Model	9.5%	0.41	10.5%	0.45
Resnet MSE Loss	9.2%	0.43	10.1%	0.51
Resnet Quantile Loss	7.8%	0.48	9.8%	0.59
Tree-Base Model + ResNet Quantile Loss	6.9%	0.5	9.7%	0.6

Table 3. The prediction Metrics of Ion and Vth.

Electrical properties	MAPE	R^2	Accuracy within 2 σ
Ion	6.9%	0.5	95.0%
Vth	9.7%	0.6	95.8%

3.3. Interpretability Notes

For the random forest tree model and resnet model, the Shap value is used to analyze the interpretability of the key characteristic parameters. Figure 6 shows the top10 important characteristic parameters that affect the electrical value. shap value can be used to assist in the positioning analysis of parameter factors that lead to poor electrical performance when exceedance and early warning occur.

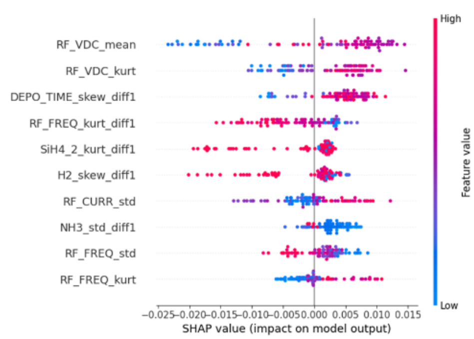


Figure 6. Shap Value.

4. Application

Based on the prediction ability of random forest and residual neural network algorithm, the model is deployed and run in the manufacturing system, and the machine production data from the FDC system is received. The prediction model and results are stored and MFA big data analysis system is used to combine the prediction results with SPC (statistical process control system) to realize virtual measurement and monitoring. The prediction model and MFA are equipped with automatic update/active learning functions, full data re-training, wide learning range, and timely learning and feedback for anomalies. At present, the MFA of the model has been deployed online, and the system is running stably. During the operation of the prediction model, the SPC quality system has predicted ten thousand pieces of glass, and the number of abnormal predicted and measured samples is 6Pcs.

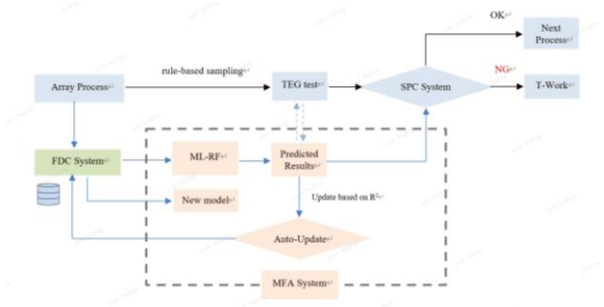


Figure 7. Flowchart of system deployment.

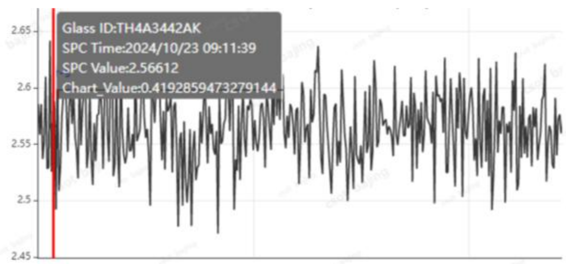


Figure 8. Predictive Time Series Diagram of the MFA and SPC system.

After the production data collection is completed, a timing task is triggered, and each glass is predicted 100%. Based on the predicted single point value, the mean value of the large plate is calculated. Combined with the low-frequency sampling inspection mechanism of the production line, the measured value of the glass can be obtained synchronously, and the model ability can be monitored based on the measured and predicted data.

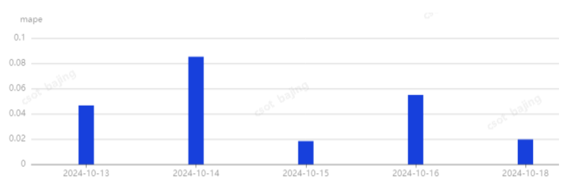


Figure 9. Predictive model indicators of the MFA system.

After the MFA system is deployed and packaged through Docker and Kafka tools, the predictive model image continues to run in the background. As can be seen from the system model kanban section in Figure 9, the MAPE is stable within 10% during the operation.

5. Conclusion

In this paper, a virtual measurement method based on the integration of random forest and deep residual neural network is proposed, which can predict the on-state current Ion & threshold voltage Vth of the display panel in real time. The prediction accuracy of the local modeling model is more than 95%, the MAPE is less than 10%, and the R^2 is between 0.52 and 0.6. The MAPE of the model stable operation is less than 10%, and the virtual measurements are linked to the SPC system to achieve accurate interception of anomalies. The automatic learning of the algorithmand system development configuration is more capable, which can ensure the prediction accuracy of the model in the process of process drift and step function change caused by preventive maintenance interference, and realize autonomy. When batch prediction anomalies occur, the factors leading to anomalies can be analyzed based on interpretability analysis to assist business personnel to locate poor results. The idea of this method can be used in the subsequent prediction of panel characteristic values based on key production parameters, and a 100% prediction monitoring model can be establishe.

6. Acknowledgements

The authors would like to thank the team at Getech Co., Ltd. for their support in implementing the ML model into our LCD production lines.

7. References

1. A.C. Diebold. “Overview of metrology requirements based on the 1994 National Technology Roadmap for semiconductor”. Advanced Semiconductor Manufacturing Conference and Workshop 1995. ASMC 95 Proceedings. IEEE/SEMI 1995, November 1995, pp. 50–60.
2. A.A. Khan, J.R. Moyne and D.M. Tilbury. “An Approach for factory-wide control utilizing virtual metrology”. IEEE Transactions on semiconductor Manufacturing, n. 20 vol.4, November 2007, pp. 364–375.
3. N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer. “SMOTE: Synthetic Minority Over-sampling Technique”. Journal Of Artificial Intelligence Research, Volume 16, pages 321-357, 2002.
4. Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001).
5. He K , Zhang X , Ren S ,et al.Deep Residual Learning for Image Recognition[J].IEEE, 2016.DOI:10.1109/CVPR.2016.90.
6. Lundberg, Scott M. and Su-In Lee. “A Unified Approach to Interpreting Model Predictions.” Neural Information Processing Systems (2017).
7. Narine Kokhlikyan, Vivek Miglani. “Captum: A unified and generic model interpretability library for PyTorch.” Neural Information Processing Systems (2017).