Convolutional neural network (CNNs) based image diagnosis for failure analysis of power devices

Akihiko Watanabe[a,*], Naoto Hirose[b], Hyoungseop Kim[b] and Ichiro Omura[c]

1 line space

[a] *Department of Electrical Engineering and Electronics, Kyushu Institute of Technology, Kitakyushu, Japan*
[b] *Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Kitakyushu, Japan*
[c] *Department of Biological Functions Engineering, Kyushu Institute of Technology, Kitakyushu, Japan*

Abstract

An image diagnosis by deep learning was applied to failure analysis of power devices. A series of images during a process to failure by power cycling test was used for this method. The images were obtained by a scanning acoustic microscopy of our real-time monitoring system. An image classifier was designed based on a convolutional neural network (CNNs). A developed classifier successfully diagnosed input image into a normal device and an abnormal device. The accuracy of classification was improved by introducing a pre-training and an overlapping pooling into the system. A technique to extract a feature related a failure is essential for the failure analysis based on the real-time monitoring and the deep learning is one likely candidate for it.

1. Deep Learning applied for failure analysis of power devices

The development of power devices is remarkable and further market expansion in the future is expected. Moreover, devices using wide band gap semiconductors are put on the market and new packaging technology for such next-generation devices are developed rapidly. Since these devices expected to use under high frequency, high power density and harsh condition, new technique is required to ensure the reliability of them.

We have proposed a failure analysis based on a real-time monitoring (RTM) [1-6]. The RTM enables to visualize the mechanism that leads to a failure by obtaining the change of structure inside the device, current distribution, electromagnetic field distribution and temperature distribution in time domain with high spatial resolution. The RTM generates time-series huge image data but it is almost impossible for human senses to extract and to judge only changes related to failure from them. Therefore, together with the technique to recognize such slight change in the data automatically, the failure analysis by RTM can increases its utility.

In last decade, the development of deep learning (DL) has been remarkable [7]. Especially, a convolutional neural networks (CNNs) is utilized to an image recognition and has been successful for automatic diagnosis in the medical field [8-10].

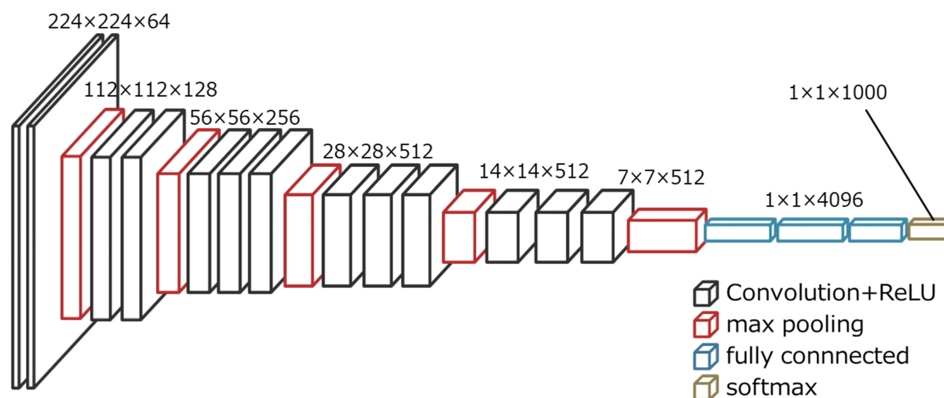

Fig. 1. A structural model of image diagnosis by CNNs.

* Corresponding author. nave@elcs.kyutech.ac.jp
Tel: +81 (93) 884 3298 ; Fax: +81 (93) 884 3298

Copyright (c) 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/

Moreover, the DL have begun to be taken in microelectronics [11-14]. Fig. 1 shows an example structural model of image diagnosis based on CNNs [15].

In this paper, we applied DL to the RTM for failure analysis of power devices. To apply DL to RTM, we developed a classifier that diagnoses a series of scanning acoustic microscopy (SAM) image obtained by RTM into "normal device" and "abnormal device". Moreover, we proposed the method to figure out a mechanism of failure mode by using RTM data with a deep learning.

2. CNNs based image diagnosis

2.1. Studied failure mechanism

To apply DL to RTM, we validated a classifier that judges a series of scanning acoustic microscopy (SAM) image obtained by RTM into two classes, "normal" and "abnormal". We used a set of images obtained when a failure had occurred caused by a degradation of wire bonding during power cycling test [1]. Fig. 2 shows the representative SAM images for training and testing of the classifier. The real-time monitoring was performed with a diode on a DBC substrate. The bonding condition was observed by SAM from the DBC substrate side. A failure occurred due to bond wire lift-off after 4677 power cycles. In this case, as the SAM image was observed each 100 cycles, we obtained 48 images adding initial and failed image by the power cycling test.

2.2. Data set preparing

We assigned an area of interest (AOI) around the diode and extracted the area of 100 × 100 pixels form original 300 × 300 pixels image. To prepare a data set for the training and testing the classifier, we have to classified the images to "normal" device and "abnormal" device. As shown Fig. 2, it is almost impossible to distinguish a difference among these images by a human vision. A behavior of forward voltage V_F of the DUT at the power cycling test was used as a criterion. A remarkable change of V_F was observed after 4300 cycles and in a few ten cycles just before the failure occurred. Moreover, an image change at bonding position was detected in subtract images after 4400 cycle [1]. Therefore, we categorized the 4 images after 4400 cycles as "abnormal device".

The number of images of 48 is too small for the training and testing the classifier. Therefore, we generated 3 images by 90°-, 180°-, 270°-rotated and 4 mirrored images from the 48 images, consequently, we prepared a data set of 384 images.

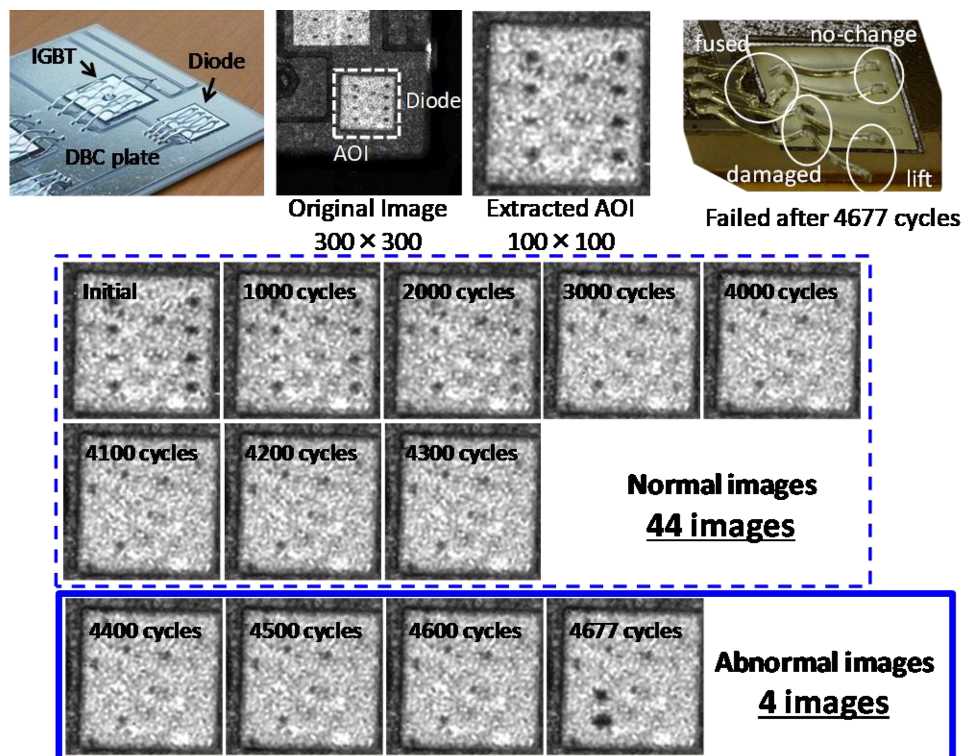

Fig. 2. SAM images of a process to failure by poser cycling test.

Table 1
Evaluation criterions

	Normal device	Abnormal device
Normal label	*a* (correct)	*c*
Abnormal label	*b*	*d* (correct)

2.3. CNN model

In this study, CNNs was applied to classification of the SAM image. In medical field, it is possible to design highly accurate classifiers by CNNs, because useful methods of feature extraction are established by using a dictionaries of medical images for learning. However, there is no precedent and image sample for learning is very few in our case. Therefore, we had to select appropriate CNN model and modify it to apply this method to SAM images of power device.

The classifier was designed based on a model VGGNet [15]. Fig. 1 shows a conceptual diagram of VGG16 model which constructed with 13 convolutional layers and 3 fully connected layer. We modified this model by a fine tuning and an overlapping pooling [7, 16]. The fine tuning is a process of learning the SAM images for a pre-training model. We applied ImageNet [17] for the pre-training. The input weights determined by pre-training are frozen so that weight parameter does not update during learning. The optimum number of the layer to freeze was experimentally obtained 14th. In the overlapping pooling, a pixel area is overlapped when an aggregate operation by a pooling. This method can improve the accuracy of aggregate by suppressing an over learning that is likely to occur with a small data set like this case.

To prevent a change of weight parameter obtained the fine tuning during training, the weight is frozen up at optimal layer. The optimal layer was obtained experimentally by using three models of the pre-trained weight was frozen to 18th layer, 15th layer and 11th layer. When a same data set was used for training and testing of these models, the highest accuracy was obtained when the weight was frozen to 15th layer.

Fig. 3 shows a structure of our proposed model. In this model, the pre-trained weights are frozen up at the 14th layer, and overlapping pooling is applied to the pooling layer of 15th and 19th layer. The detail of this model is shown in Appendix.

2.4. Evaluation result

The classifier was evaluated by a *k*-fold cross-validation [7]. The data set of the SAM image was randomly divided into four data sets. Each data set constructed by 96 images including 88 images of “normal” device and 8 images of “abnormal” device. The three data sets were used for training and the original images without augmentation data was used for testing. A result of the classification was categorized as shown in Table 1 and CCN model was evaluated with the following value:

$$accuracy = (a + d) / (a + b + c + d) \quad (1)$$

$$True\ Positive\ rate\ (TP) = a / (a + b) \quad (2)$$

$$False\ Positive\ rate\ (FP) = c / (c + d) \quad (3)$$

As for ideal classifier, the value of accuracy, TP and FP is 1, 1 and 0 respectively. In addition, the value of TP and FP indicates accuracy of judgement for “normal” device and “abnormal” device respectively. These values obtained by testing were different for each test because of a bias of data set, therefore we used an average value which obtained by 5 tests of each data set.

We compared three CNN models constructed

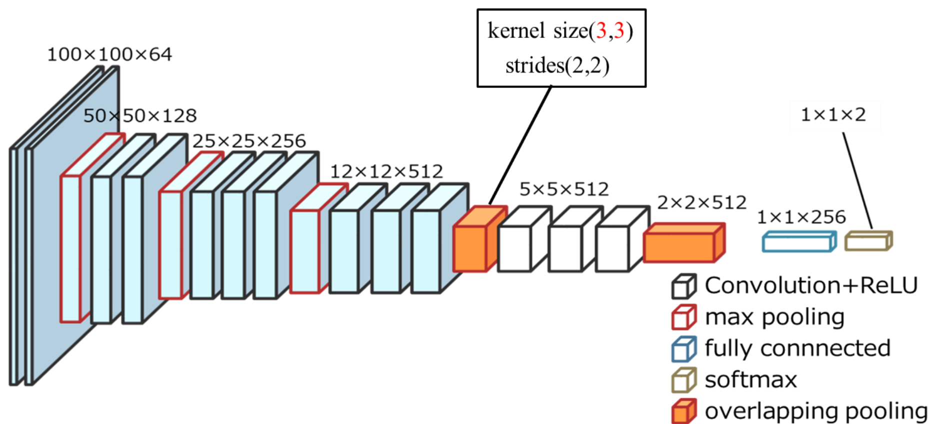

Fig. 3. Proposed model of image diagnosis by CNNs.

Table 2
Result of 4-fold cross validation

	accuracy	*TP*	*FP*
without Pre-training	0.9225	1.0000	0.8938
with Pre-training	0.9406	0.9659	0.3375
Pre-training + Overlapping pooling	0.9714	0.9818	0.1438

based on the VGG16. One is a model without frozen up of the pre-trained weight. Second is a model pre-trained weight are frozen up at the 14th layer, and our proposed model which the pre-trained weights are frozen up at the 14th layer and overlapping pooling is applied to the pooling layer of 15th and 19th layer. The results are shown in Table 2. If a CNN model judges all images in the data set as "normal" device, the accuracy is 0.92, TP = 1.0 and FP = 1.0 (a = 88, b = 0, c =8, d = 0). The result of first model without frozen up of pre-trained weight was indicated almost same value of this case. This result means that this model could not construct an accurate classifier. In the case of second model which the pre-trained weight is frozen up, the accuracy and FP were improved. Moreover, our proposed model with overlapping pooling indicated good score in all parameter. The value of 0.14 in FP means judgement accuracy for "abnormal" device is lower. This result caused by much fewer "abnormal" device image than that of "normal" device images in the data set.

3. Advantage of failure analysis based on Real-Time Monitoring with Deep Learning

A failure mode of power devices is determined by a correlation between physical phenomena occurring during process to failure. In principle, it is possible to record all these phenomena by RTM if appropriate parameters are selected for monitoring. The RTM we have proposed monitors and records the state change of the power device during acceleration test over whole period of time. The object to be monitored includes not only to electrical characteristics and heat dissipation characteristics but also to internal structural changes *etc*. The obtained data includes whole information of a failure mechanism from occurrence of a microscopic phenomenon which is trigger of a failure to degradation process to failure. In the conventional analysis represented failure analysis, the cause of failure is determined by static observation with opening the failed device. On the other hand, in the real-time monitoring, microscopic phenomena which can be cause of failure are dynamically observed in time series before a failure occurs. Therefore, the RTM has a great advantage for extraction of the relationship between physical properties of constitutional materials and degradation phenomena inside the device or extraction of phenomena that cannot be modelled in simulation. The RTM monitoring proposes a clear spec for materials and a new evaluation method for material development, and also it opens the way to the development of a highly accurate simulation model (see Fig. 4).

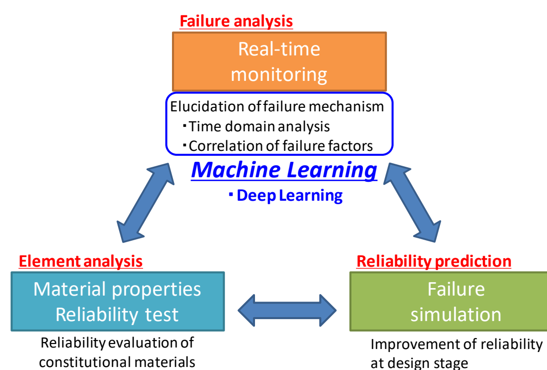

Fig. 4. Method for enhancing reliability of a power semiconductor based on a real-time monitoring.

To figure out a mechanism of failure mode by using RTM data, it is necessary to detect and quantify the parameter changes related to the failure. A change in monitored parameter is quantified and these change is indexed by a type and degree of the change. The time-series indexes systematically express a failure mechanism. By finding a pattern depending on a failure mode from the time series index, it is possible to predict future failure modes from RTM data during the acceleration test. Moreover, it is also possible to estimate a life time during accelerated test by judging the progress of deterioration from an occurrence timing of index and the degree of parameter change. If the target failure mode is limited, *e.g.* degradation of die attach or wire bonding by a power cycling test, we consider that the analysis method of a failure mechanism is applicable with monitoring of representative parameters without monitoring all parameters.

A data analysis by deep learning is suitable for a failure analysis by RTM. The most basic role of DL is extraction and judgement of a parameter change related to a failure mechanism from RTM data. The data to be learned in DL is not limited to one type but by combining several types of data, it is possible to design AI that can make more complicated judgment.

4. Conclusion

An image classifier was designed based on a convolutional neural network (CNNs) for failure analysis of power devices by a real-time monitoring. A developed classifier successfully diagnosed input SAM image into a normal device and an abnormal device. The accuracy of classification was improved by introducing a pre-training and an overlapping pooling into the system. Since the small number of data sets, we plan to new validation test with new data sets in the future.

References

[1] Watanabe A, Tsukuda M, Omura I. Failure analysis of power devices based on real-time monitoring. Microelectronics Reliability, 55 (2015) 2032-2035.

[2] Watanabe A, Omura I. A power cycling degradation inspector of power semiconductor devices. Microelectronics Reliability, 88-90 (2018) 458-461.

[3] Watanabe A, Nagao R, Omura I. Real-time imaging of temperature distribution inside a power device under a power cycling test. Microelectronics Reliability, 76-77 (2017) 490-494.

[4] Shiratuchi H, Matsushita K, Omura I, IGBT chip current imaging system by scanning local magnetic field. Microelectronics Reliability, 53 (2013) 1409-1412.

[5] Hirata N, Watanabe A, Omura I. High speed real-time temperature monitoring system inside power devices package using infrared radiation. Proc. of 2014 International Conference on Solid State Devices and Materials, 2014, 1016-1017.

[6] Tsukuda M, Matsuo K, Tomonaga H, Okoda S, Ryuzo N, Tashiro K, Omura I. Magnetic flux signal simulation with 16-channel sensor array to specify accurate IGBT current distribution. Proc. of International Conference on Integrated Power Electronics Systems, 2016, 264-268.

[7] Goodfellow I, Bengio Y, Courville A. DEEP LEARNING. The MIT Press, London, 2016.

[8] Li H *et al.* An improved deep learning approach for detection of thyroid papillity cancer in ultrasound images. Scientific report 8, 2018, Article number 6600.

[9] Şahan S, Polat K, Kodaz H, Güneş S. A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis. Computers in Biology and Medicine 37, 2007, 415-423.

[10] Pérez N *et al*. Improving the performance of machine learning classifiers for Breast Cancer diagnosis based on feature selection. Proc. of the 2014 Federated Conference on Computer Science and Information Systems, 2014, 209-217.

[11] Khera N, Khan S. Prognostics of aluminum electrolytic capacitors using artificial neural network approach. Microelectronics Reliability, 81, 2018, 328-336.

[12] Sun X, Huang M, Liu Y, Zha X. Investigation of artificial neural network algorithm based IGBT online condition monitoring. Microelectronics Reliability, 88-90, 2018, 103-106.

[13] Veenhuizen M. Void detection in solder bumps with deep learning. Microelectronics Reliability, 88-90, 2018, 315-320.

[14] Zhao L *et al*. Optimization of an Artificial Neural Network System for the Prediction of Failure Analysis Success. Microelectronics Reliability, 92, 2019, 136-142.

[15] Simoyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations, Vol. 3, 2015, 1-14.

[16] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. Proc. Advances in Neural Information Processing Systems, Vol.25, 2012, 1090−1098.

[17] ImageNet. http://www.image-net.org/index (access:2019/02/03).

Appendix

Detail of network model

Layer	Filter size	Stride	Output size	Remarks
input_1	-	-	(100,100,3)	-
block1_conv1	(3,3)	(1,1)	(100,100,64)	ReLU
block1_conv2	(3,3)	(1,1)	(100,100,64)	ReLU
block1_pool	(2,2)	(2,2)	(50,50,64)	
block2_conv1	(3,3)	(1,1)	(50,50,128)	ReLU
block2_conv2	(3,3)	(1,1)	(50,50,128)	ReLU
block2_pool	(2,2)	(2,2)	(25,25,128)	
block3_conv1	(3,3)	(1,1)	(25,25,256)	ReLU
block3_conv2	(3,3)	(1,1)	(25,25,256)	ReLU
block3_conv3	(3,3)	(2,2)	(25,25,256)	ReLU
block3_pool	(2,2)	(2,2)	(12,12,256)	
block4_conv1	(3,3)	(1,1)	(12,12,512)	ReLU
block4_conv2	(3,3)	(1,1)	(12,12,512)	ReLU
block4_conv3	(3,3)	(2,2)	(12,12,512)	ReLU
block4_olp	(3,3)	(2,2)	(6,6,512)	Overlapping
block5_conv1	(3,3)	(1,1)	(6,6,512)	ReLU
block5_conv2	(3,3)	(1,1)	(6,6,512)	ReLU
block5_conv3	(3,3)	(2,2)	(6,6,512)	ReLU
block5_olp	(3,3)	(2,2)	(3,3,512)	Overlapping
Flatten	-	-	(3,3,512)	Flatten
Dense	-	-	(1,1,256)	Dense
Dropout	-	-	(1,1,256)	Dropout
Dense	-	-	(1,1,2)	Softmax