A more reliable defect detection and performance improvement method for panel inspection based on artificial intelligence

Eui-Young Jeonga, Jaewon Kimc, Won-Hyouk Jangb, Hyun-Chang Limb, Hanaul Noha and Jong-Myong Chobic

aInspection Equipment Development Team, Samsung Display, Yongin-si, Korea; bAILab., Samsung Display, Yongin-si, Korea

ABSTRACT
This paper presents a practical approach to automatic inspection of display panels based on deep neural networks. The approach accurately detects appearance defects on display panels in various sizes and shapes within a short computation time. We propose a novel reliable detection network using the multi-channel parameter reduction method, which preserves high-resolution features of defects at sub-sampling steps of convolutional operations. Our proposed network consists of two sub-networks with different functions: pixel-wise segmentation of defect regions and distinction of real defects from fake defects. Compared with conventional deep learning networks, the proposed network achieved a more accurate detection rate, i.e. an F1-score of 81%, for real defect images acquired from an actual display manufacturing process. In addition, we propose a conditionally paired generative network that generates synthetic images of scarce defects under four different lighting conditions. The proposed networks improved the detection accuracy and can be applied to automatic inspection processes in display manufacturing factories in place of human inspection.

1. Introduction
Appearance inspection is one of the most crucial processes in the manufacture of organic light-emitting diode (OLED) displays because it is the final process in the classification of a manufactured panel into a good or flawed panel. Human inspectors traditionally performed the process, but its reliability was highly dependent on the degree of proficiency and the physical condition (such as the wellness or unwellness) of the inspector. Also, some defects were quite challenging for human inspectors to detect within a short period of time. To overcome these limitations associated with human inspectors, automatic inspection techniques have been applied in the display manufacturing industry. Such techniques generally use an image processing algorithm to inspect display panels automatically by analyzing their visual information. Traditional inspection algorithms performed well in detecting pre-defined features of defects but not in detecting undefined features. They also posed a challenge in setting numerical criteria for distinguishing real defects from fake or allowable defects. To overcome the limitations of conventional inspection algorithms, deep learning (DL) methods [1,2] have recently been widely explored. For a DL-based inspection algorithm to be a successful alternative to conventional algorithms, two key requirements must be satisfied. First, it is crucial to detect defects in various locations, sizes, and shapes by processing high-resolution images within a limited time. Second, the detection rate of real defects must be maximized while minimizing the false detection rate.

In this paper, we propose a novel deep neural network for robust detection of real defects on display panels that outperforms well-known DL methods with respect to the two aforementioned requirements. In addition, a conditionally paired generative network is proposed to enhance the detection rate by synthetically increasing feature domains in a given image database. Finally, we validate the performance of our proposed networks by evaluating their defect detection rate using images of display panels that were acquired from an actual manufacturing process.

2. Prior studies
2.1. Deep neural network-based object detection models
Object detection methods based on DL can be classified into two categories: patch-based detection methods and pixel-based segmentation methods.
The patch-based methods generally sample pixels in a specific region of interest (ROI) and process them sequentially by scanning them over the whole image. The probability of the target objects being included is calculated in each ROI, and the ROI with the highest probability is chosen as the target region. Such a sliding window operation may be suitable for low-resolution input images but not for high-resolution images because it takes several tens of seconds to calculate and discriminate hundreds to thousands of ROIs. Region proposal networks such as Faster R-CNN [3] and Mask R-CNN [4] were recently proposed. Regression networks such as YOLO [5] and SSD [6] achieve real-time computation but have limitations in dealing with objects in various sizes because the regression networks have a fixed anchor size.

The pixel-based segmentation algorithm uses a fully convolutional neural network (FCN), which optimizes feature extraction and dimensionality reduction through multi-layer operations and identifies an object region in a pixel unit, as shown in Figure 1. The FCN models achieve fast computational speed, comparable to those of the regression network models. Thus, pixel-based segmentation using FCN models is suitable for fast detection tasks, such as inspection of defects in manufacturing processes, and for target objects in various sizes, locations, and shapes.

Long et al. [7] proposed an FCN-based segmentation algorithm that uses skip connection layers to compensate for the location information of features lost in the pooling operations. Semantic segmentation [8] tackled the problem of the rapid expansion of the image size by 32, 16, and 8 times. Afterwards, SegNet [9] improved the segmentation prediction by using the advantages of the deconvolution network and supplementing the location information of the feature by pooling location map information.

In this paper, we propose a novel FCN algorithm for pixel-based segmentation that inherits the advantages of the previous methods of detecting objects in various locations, sizes, and shapes. By designing a lightweight network comparable to those in the previous methods, we minimized the learning and inference times.

2.2. Deep neural network-based generative models

Goodfellow et al. [10] proposed a generative model based on a deep neural network. In the model, network parameters were trained to optimize the adversarial objective function of a discriminator and a generator. Since then, generative neural network models have been widely researched, including conditional generative adversarial networks (conditional GANs) [11], which trained a model by adding a conditional label as an input and provided a basis for the control of randomly generated data. Deep convolutional GAN [12] showed that synthetic data can be controllably generated via vector operation over the image features. The structure and style of the image database can be learned by S2-GAN [13], after which an input image can be transformed into a synthetic image in the same style as that of the images in the image database using the learned model.

Cycle-GAN [14] converts an original style to other styles while maintaining the structure of the image. The method could be applied to the generation of synthetic defect images by inputting a normal image without defects.

Summarizing the previous deep generative model approaches, the distribution of the latent variable Z that represents data X is assumed as a normal distribution and optimized through feature extraction by a discriminator. They also generated a single-style low-resolution (28 × 28 ∼ 128 × 128) image. However, an inspection process for defects on display panels requires high-resolution images, as shown in Figure 2, under various illumination conditions (Due to the company’s security policy, we cannot specify these conditions and for what purposes we used four different lighting conditions) to maximize the visibility of the defects depending on their shape, location, physical characteristics, and other attributes. Figure 2 (b) shows one defect in different illumination conditions.

3. Proposed neural network models

3.1. Multi-channel parameter reduction and reliable detection network

Our proposed inspection system takes multiple 5,120 × 9,500-pixel photographs of a display panel under different illumination conditions, which were designed to maximize the visibility of various kinds of defects. For mass production lines to be efficient, the inspection process for a display panel must be completed within 1
second. As a deep neural network is generally associated with hundreds of thousands of computational parameters that directly affect the inference speed of the network, it is crucial to minimize the number of computational parameters to meet the required inspection speed. In this study, we propose a novel multi-channel algorithm for reducing the size of the input parameters while maintaining the key feature information of a defect shape.

The proposed method first samples pixels from given multi-channel images at 4-pixel intervals, as shown in Figure 3, which reduces the number of feature parameters by one-fourth compared to using the original images. Figure 4 compares our proposed sampling method (e) with four conventional resizing methods (a-d), with our proposed method showing the feature shape closest to that of the original image.

As shown in the image in Figure 5, the pixel-based segmentation algorithm could not distinguish between true defects (b) and false defects (a), such as contaminants, which resulted in a high false detection rate. Therefore, we propose a robust detection method based on DL to minimize false detections.

We propose a reliable detection network that uses the multi-channel parameter reduction method. This method uses, as inputs, four photographs taken under different lighting conditions. Our learning architecture in Figure 6 consists of two types of neural network: a pixel-based segmentation network for locating defects and a classification network for determining whether the candidate defect regions are real or fake. In this study, the proposed network utilized a conventional network to extract the features of the input images and designed lightweight upsampling layers to maximize the inference speed.

Equation (1) is the loss function of the segmentation and classification network. Our proposed network learns to minimize the total cost of Equation (2). Weights $\alpha$ and $\beta$ in Equation (2) are user-defined values for segmentation and classification performance control.

$$\frac{1}{N} \sum_{i} \{Y \log \hat{Y} + (1 - Y) \log(1 - \hat{Y})\}$$ (1)

$$\arg \min_w \{\alpha L_{\text{segmentation}} + \beta L_{\text{classification}} + L_2(w; e^{-4})\}$$ (2)

### 3.2. Conditionally paired generative neural network

In previous studies on generative networks, the initial value of the latent variable $Z$ that represented data $X$ was defined as the value of one of 100–1,000 random variables with normal or uniform distributions, and the optimal size of the latent variable was evaluated experimentally. For generating high-resolution images, however, it is impossible to evaluate the optimal size of the latent variable. To estimate the initial size of the latent variable, a convolutional autoencoder [15] could be used. In this method, an encoder learns features, $Z$, that represents data $X$, and a decoder generates data $X$ from $Z$, as shown in Figure 7. This method is useful in defining low-dimensional optimal features or latent variables to represent data $X$. To accelerate the learning speed of our conditionally paired generative network, we converted a $512 \times 512$-pixel image into a $512$-dimensional feature

---

**Figure 2.** Defect images used for appearance inspection of display panels. (a) Three types of defects in an optimized lighting condition. (b) One defect in three different lighting conditions.

**Figure 3.** The proposed multi-channel parameter reduction method.
Figure 4. Result of the application of the image reduction algorithm. The proposed method resulted in the shape that was most similar to that of the original image.

Figure 5. Comparison of images of false defects (contaminants) and true defects.

Figure 6. The proposing deep learning architecture for accomplishing robust detection performance.

As shown in Figure 2, our inspection system takes four photographs of a display panel under different illumination conditions, assuming that the visibility of the defects on the panel can be maximized by one of the conditions. We designed the novel conditionally paired generative network shown in Figure 8 to generate 4-channel images with the same style as that of the obtained photographs.

As described above, our conditionally paired generative network utilizes the initial value of the latent variable $Z$, which the convolutional autoencoder optimized. Our network can generate a large variety of defect images in terms of position, size, and shape under four different lighting conditions by adjusting the control parameters. Equations (3) and (4) are the objective functions of the vector using the convolutional autoencoder in Figure 7 and global average pooling, and used the result as the initial latent variable space.
Figure 7. The convolutional autoencoder for defining the initial latent variable $Z$ for the proposed conditionally paired generative network.

Figure 8. Our conditionally paired generative network.

discriminator and the generator in our network, respectively. It produces synthetic images that are closest to the obtained photographs through competition between the generator for making synthetic images and the discriminator for judging if an image is real or fake. The parameters of a generator network are trained to minimize the loss defined by the discriminator.

$$
\arg\min_w \left\{-\frac{1}{N} \sum \{ Y \log \hat{Y} + (1 - Y) \log (1 - \hat{Y}) \} + L_2^w(e^{-4}) \right\}$$

(3)

$$
\arg\min_w \left\{ \sum D_l(G_i(Z))_{\text{Fake}} + L_2^w(e^{-4}) \right\}
$$

(4)

When the discriminator classifies the four images generated under different lighting conditions as real images, the single set becomes a candidate for learning data and this process is repeated until enough learning data are generated.

4. Experimental evaluation

4.1. Experiment dataset and environment

We used Tensorflow v1.13 [16] and Keras v2.2 [17] to implement our reliable detection network and conditionally paired generative network. A single GPU, P100, was used for the training. To validate the performance of our method, we tested the chipping defect data on a wide range of locations and shapes, as shown in Figure 9, which we obtained from our inspection system for actual manufacturing lines of display panels. The proposed models are applicable to all defects. However, due to the company’s security restrictions, the evaluation results are shown only for the chipping defect of the window glass, which is an important defect that causes great loss because all the assembled panels and accessory materials cannot be used. For accuracy evaluation, the Precision and Recall scores [18] were calculated for our model and conventional models.

4.2. Evaluation and analysis

4.2.1. Comparison of detection performance: F1-score and inference speed

To verify the accuracy of our detection network, it was evaluated in comparison with U-Net [19], YOLO v2 [20], well-known object detection models, and commercial tools such as V-Tool [21] and S-Tool [22]. Table 1 compares the F1-scores of the methods for the test dataset that consisted of 30 actual chipping defects and 150 false defects, i.e. contaminants. A learning dataset with 100 actual chipping defects was used for each method.
Figure 9. Sample images of chipping defects in various locations and shapes.

Table 1. Results of the comparison of the defect detection methods.

| Net. Model | True Positive | False Positive | F1-score | Inference Time (@GTX1070) |
|------------|---------------|----------------|----------|--------------------------|
| Proposed network | 76.7% | 2.7% | 81% | 0.47 second/cell |
| U-Net | 80.0% | 46.7% | 45% | 0.45 second/cell |
| YOLO (ver. 2) | 70.0% | 64.0% | 35% | 0.48 second/cell |
| V-Tool | 73.3% | 52.0% | 40% | 0.96 second/cell |
| S-Tool | 83.3% | 8.0% | 75% | 0.71 second/cell |

Our network outperformed U-Net, YOLO, and the commercial tools, with an 81% F1-score. The inference speed of all the methods was faster than 1 second/cell, which is the speed limit for the inspection process in a display manufacturing line. Our network accomplished 1.5 times faster computing than the commercial tools. S-Tool showed the highest true positive rate of 83.3%. However, its false positive rate was 2.9 times higher than that of our network, which can degrade the efficiency of an automatic inspection process by requiring re-inspection by humans. Overall, our method achieved a true positive rate that was comparable with that of the best method, S-Tool, because the model with the highest accuracy of classification was selected as the best our model, but a false positive rate, 2.7%, that was significantly lower than the false positive rate of S-Tool, 8.0%, which was the second lowest rate.

4.2.2. Improvement of detection performance using the generative network

Synthetic defect datasets can increase the accuracy of a detection network while preventing the underfitting and overfitting caused by insufficient learning data. Our conditionally paired generative network generates synthetic defect images that are very similar to photographs of actual defects, as shown in Figure 10. The labeled defect regions in the actual photographs are converted into the latent variable z, and synthetic defect images that correspond to z are produced in our generative network. As a result, the defect database can be significantly extended to include new defect images in various positions and with various shapes.

Figure 11 shows the chipping defect images generated under different lighting conditions from sample z. The discriminator in our network classified the defect images into real defect photographs, so there was little difference between them and the actual defect photographs acquired by our inspection system.

Figure 12 shows bad generation results, which the discriminator adjudged as photos of partially fake defects. Compared to the images in Figure 11, the resulting fake defect photographs had irregular and abnormal shapes, which were not observed in the actual defect photographs.

We performed a two-dimensional (2D) principal component analysis (PCA) of the actual photographs and the synthetic images to compare their feature distributions. The results are shown in Figure 13. In the synthetic images, the distribution of the real images, i.e. the green dots, is clearly distinguished from the distribution of the fake images, i.e. the red dots. In contrast, the real images have the same distribution as the real photographs, i.e. blue dots, which means they are visually identical.

To validate the improvement of the detection performance, we selected various types of real images to extend the given learning database. The test results are summarized in Table 2. The true positive rate of our detection network improved by 10% from 76.7% to 86.7% with the use of the extended database that included synthetic defect images. There was no significant difference between the false positive rate when the original dataset was used (2.7%) and when the extended dataset was used (3.3%). Therefore, our conditionally paired generative model effectively improved the performance of the detection network.

As the amount of learning data increased from 100 to 150 and 200 cells, the detection performance gradually improved from 76.7% to 80% and 86.7%. We attribute the improvement of the detection performance to the addition of the synthetic defect images to the learning data that had features similar to those of the actual defects in the test database, as shown in Figure 14. The accuracy of a neural network is closely related to the generality of the learning database. To achieve robust performance, sufficient evaluation and learning data must be
Figure 10. Our generative network produces new defect images by combining defect regions with normal regions.

Figure 11. Results of the defect generation corresponding to the four lighting conditions.
Figure 12. Comparison of the real and fake defect images classified by the discriminator in our generative network.

Figure 13. 2D PCA plot of the actual and generated images. The blue dots labeled ‘Real’ denote the actual photographs acquired in a visual inspection system. The green dots labeled ‘GenReal’ were classified as real images by the discriminator in our network. The red dots labeled ‘GenFake’ were classified as fake images by the discriminator.
collected. However, collecting sufficient data is not an easy task. Therefore, we expect our generative network to be effectively used to improve the performance of any deep learning model while increasing the generality of a given database.

5. Conclusion

In this paper, we proposed a reliable network for detecting defects of display panels in various locations and with various shapes. Our method achieved high detection accuracy with fast computing speed over high-resolution photographs obtained from a visual inspection system. To process high-resolution images within a limited time, a multi-channel parameter reduction algorithm was used to maintain the information on the defect features while reducing the image parameters. Our network achieved superior performance with an 81% F1-score compared to U-Net, YOLO, and commercial tools. In addition, we proposed a novel conditionally paired generative network for the production of synthetic defect images that reflect various lighting conditions. Our method differs from previous networks for the generator and discriminator at several architecture. Our model creates images with various styles and with a high resolution, unlike those of existing generation models that create single-style and low-resolution images. Also, our model is capable of fast and stable learning, unlike existing GAN models, using the autoencoder-based latent variable space. As a result, our method improved the true positive rate for the detection of chipping defects by 10%, from 76.7% to 86.7%. Our networks will enhance the capability of automated inspection technology and complement manual inspection operations by improving the overall reliability and efficiency of an inspection system.

In the future, it is necessary to develop a self-learning system that continuously learns generated data using quantitative evaluation methods and studies for the optimal neural architectures, such as Neural Architecture Search (NAS) [23]. We are convinced that this system will maximize the performance of artificial intelligence inspection systems.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Notes on contributors

Eui-Young Jeong received his M.S. degree, with specialization in Software Engineering, from KOREA University. He is currently working as a staff engineer at the Inspection Equipment Development Team of Samsung Display.

Jaewon Kim received his Ph.D. degree, with specialization in Computer Science, from Imperial College London. His research interests span computer vision/graphics, computational photography, and optics. He is currently working as a Vice President at the Mechatronics Technology Center of Samsung Display.

Won-Hyouk Jang received his Ph.D. degree, with specialization in Chemical Engineering, from the Texas A&M University. He is currently leading AI/ML technologies as Master (VP specialized in R&D) at the AI Lab. of Samsung Display.
Hyun-Chang Lim received his Ph.D. degree, with specialization in Analytical Chemistry, from Seoul National University. He is currently working as a principal engineer at the Mechatronics Technology Center of Samsung Display.

Hanaul Noh received his Ph.D. degree, with specialization in Nanoscale Physics, from Seoul National University. He is currently working as a principal engineer at the Mechatronics Technology Center of Samsung Display.

Jong-Myong Choi received his Ph.D. degree, with specialization in Industrial Engineering, from Iowa State University. He is currently working as a principal engineer at the AI Lab. of Samsung Display.

References
[1] B. Staar, B. Staara, M. Lütjena, and M. Freitag, Anomaly detection with convolutional neural networks for industrial surface inspection, Procedia CIRP 79 (1), 484–489 (2019).
[2] C. C. Huang, et al., Study on Machine Learning Based Intelligent Defect Detection System. MATEC Web of Conferences. EDP Sciences 01010, 2018.
[3] S. Ren, et al., Faster r-cnn: Towards real-time object detection with region proposal networks. Adv. Neural Inform. Process. Syst. 39 (6), 91–99 (2015).
[4] K. He, et al., Mask r-cnn. Proceedings of the IEEE international conference on computer vision. 2017, pp. 2961–2969.
[5] J. Redmon, et al., You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, pp. 779–788.
[6] W. Liu, et al., SSD: Single shot multibox detector. European conference on computer vision. Springer, Cham, 2016, pp. 21–37.
[7] J. Long, E. Shelhamer, and T. Darrell, Fully convolutional networks for semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition. 2015, pp. 3431–3440.
[8] H. Noh, S. Hong, and B. Han, Learning deconvolution network for semantic segmentation. Proceedings of the IEEE international conference on computer vision. 2015, pp. 1520–1528.
[9] V. Badrinarayanan, A. Kendall, and R. Cipolla, Segnet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE T. Pattern Anal. 39 (12), 2481–2495 (2017).
[10] I. Goodfellow, et al., Generative adversarial nets. Adv. Neural Inform. Process. Syst. 27, 2672–2680 (2014).
[11] M. Mirza and S. Osindero, Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.
[12] A. Radford, L. Metz, and S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.
[13] X. Wang and A. Gupta, Generative image modeling using style and structure adversarial networks. European Conference on Computer Vision. Springer, Cham, 2016, pp. 318–335.
[14] J-Y. Zhu, et al., Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE international conference on computer vision. 2017, pp. 2223–2232.
[15] Y. Zhang, A Better Autoencoder for Image: Convolutional Autoencoder. ICONIP17-DCEC. 2017.
[16] TensorFlow, https://www.tensorflow.org/
[17] Keras Documentation, https://keras.io/
[18] Precision and recall, https://en.wikipedia.org/wiki/Precision_and_recall
[19] O. Ronneberger, P. Fischer, and T. Brox, U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015, pp. 234–241.
[20] J. Redmon and A. Farhadi, YOLO9000: better, faster, stronger. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7263–7271.
[21] Cognex Corporation, VISIONPRO VIDI, https://www.socognex.com/products/deep-learning/visionpro-vidi (2020).
[22] Saige-Research, SAIGE-VISION, 2020, http://www.saigere search.ai/home/sub.php?menukey = 61.
[23] C. Liu, et al., Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019, pp. 82–92.