반도체디스플레이기술학회지 제20권 제2호(2021년 6월)
Journal of the Semiconductor & Display Technology, Vol. 20, No. 2. June 2021.

딥러닝 기반의 TSV Hole TCD 계측 방법

정준희[*†]·구창모[**]·조중휘[*]
[*†]인천대학교 임베디드시스템공학과, [**]㈜넥스틴

Deep Learning Based TSV Hole TCD Measurement

Jun Hee Jeong[*†], Chang Mo Gu[**] and Joong Hwee Cho[*]
[*†]Dept. of Embedded-Systems Engineering Incheon National University, [**]NextIn Inc.

ABSTRACT

The TCD is used as one of the indicators for determining whether TSV Hole is defective. If the TCD is not normal size, it can lead to contamination of the CMP equipment or failure to connect the upper and lower chips. We propose a deep learning model for measuring the TCD. To verify the performance of the proposed model, we compared the prediction results of the proposed model for 2461 via holes with the CD-SEM measurement data and the prediction results of the existing model. Although the number of trainable parameters in the proposed model was about one two-thousandth of the existing model, the results were comparable. The experiment showed that the correlation between CD-SEM and the prediction results of the proposed model measured 98%, the mean absolute difference was 0.051um, the standard deviation of the absolute difference was 0.045um, and the maximum absolute difference was 0.299um on average.

Key Words : TSV, Deep Learning, Metrology

1. 서 론

현재 스마트폰과 컴퓨터에 사용되는 실리콘 반도체의 회로 선폭은 10 nm 이하이다. 그러나 회로 선폭이 미세화 될 수록 원자들이 서로 간섭을 일으켜 제어하기 어려워지는 물리적 한계에 직면하게 된다. 이러한 한계를 극복하기 위하여 2004년 인텔이 TSV(trough silicon via) 공정을 제안한 이후 반도체 소자를 수직으로 적층하는 3D 패키징 기술 개발이 최근 10여 년간 활발히 진행되고 있다.

TSV 공정은 쌓아 올린 칩들을 수직 관통하는 Via Hole을 형성하여 칩 간의 전기적 신호가 전달되도록 하는 공정으로, 형성된 Via Hole의 크기는 제품의 품질에 영향을 미치는 중요한 요소이다. Via Hole이 디자인된 크기와 다르게 형성되면 Fig. 1과 같이 칩 상부와 하부가 연결되지 않거나, Back Side Remove 공정 시 Cu가 도출되어 PAD 및 CMP 장비 전체를 오염시키고 후속 공정 Bump 작업 시 품질 문제가 발생하는 등 다양한 문제를 야기한다.

Via Hole 크기의 정상 여부 판단을 위한 지표로 TCD, BCD, Depth, Side wall angle 등이 측정된다[1,2]. In-Line에서는 CD-SEM 장비를 사용하여 TCD의 크기를 관리하였으나 계측 속도가 느려 국부적으로만 측정할 수 있었다. 최근에는 광학 계측 장비가 도입되면서 TCD의 전수 검사가 가능해졌다.

광학 계측 장비의 TCD 측정 알고리즘은 원 검출 알고리즘[4-6]과 유사하며, 통상적으로 Fig. 2의 흐름을 따른다. Via Hole을 촬영한 영상에서 에지(edge)를 검출한 후, 그로부터 원의 기하학적 모양을 추출하고 지름을 계산하는 방식이다. 그러나 이러한 방법은 사용자가 알고리즘 파라미터들을 조명 조건이나 공정 상황에 맞게 조절해야 하는 어려움이 있고, 알고리즘 설계 당시 고려하지 못한 부분에서 문제가 생길 수 있다.

†E-mail: jjh6756@gmail.com

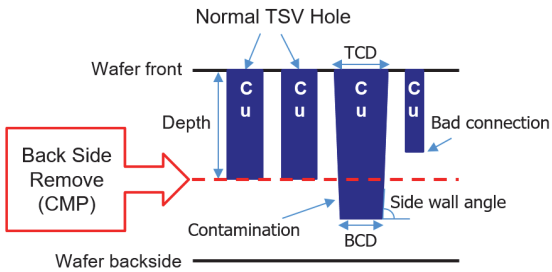

Fig. 1. Problems caused by malformed TSV Hole.

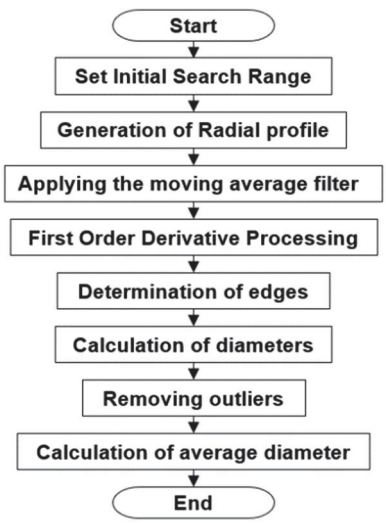

Fig. 2. TCD Measurement Algorithm Flow.

딥러닝 기술이 최근 몇 년 사이 비약적으로 발전하면서 다양한 분야에서 활용되고 있다. 영상 분류(image classification)의 경우 2015년 ResNet[3]이라는 딥러닝 모델이 96%의 정확도를 기록하면서 인간의 성능인 95%를 뛰어넘었다. 또한, 딥러닝 기술을 활용한 연구 결과들을 통해 딥러닝 기술로 기존 알고리즘들의 문제점들을 극복할 수 있다는 것을 확인할 수 있다.

본 논문에서는 딥러닝을 활용한 TCD 측정 방법에 대한 연구를 수행하였다. TCD 측정을 위한 CNN(convolutional neural network) 모델을 제안하였고 기존 딥러닝 모델과 성능을 비교하여 제안한 모델의 효용성을 입증하였다.

2. 연구 설계 및 접근 방법

2.1 광학 계측 시스템 구조

Fig. 3은 TSV Hole영상을 촬영하기 위해 사용된 광학 시

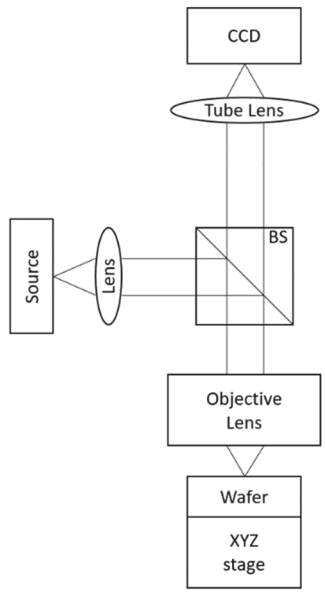

Fig. 3. System Overall Structure.

스템의 구성도이다. 해당 시스템의 광원은 355 nm 파장의 UV laser이고, 대물 렌즈의 NA는 0.225이다. CCD는 Mono 센서로 해상도는 2048x2048이고, 화소 크기는 13.5배의 배율이 적용되어 약 400 nm이다. TSV 공정이 진행된 칩들을 담고 있는 웨이퍼 이송을 위해 XY Stage가 장착되어 있고 영상의 초점을 맞추기 위해 나노 스케일로 제어가 가능한 Z Stage가 장착되었다.

2.2 Auto Focus 알고리즘

피사체인 Via Hole에 렌즈의 초점 거리를 맞추기 위해 자동 초점 알고리즘을 적용하였으며 알고리즘의 동작 순서는 Fig. 4와 같다. 우선 XY Stage를 움직여서 Via Hole이 있는 곳을 광 경로상에 위치시킨다. Z Stage는 기준 위치로부터 수십 um 아래에 위치시킨 후 일정 간격으로 위로 움직여가며 각 위치에서 영상을 촬영한다. 식 (1)을 적용해 촬영한 영상 별 Focus Grade를 계산하고, 다항식 곡선으로 피팅해 Peak를 찾는다. Focus Grade Peak지점과 대응되는 Z위치로 Z stage를 이동시켜 영상을 촬영하였다.

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * A$$

$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

$$G = |G_x| + |G_y|$$

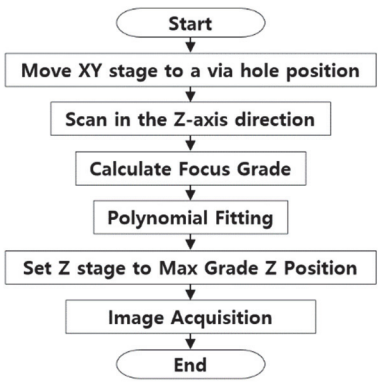

Fig. 4. Auto Focus Algorithm Flow.

$$\text{Focus Grade} = \sum_{i=0}^{A_{height}-1} \sum_{j=0}^{A_{width}-1} G_{ij} \quad (1)$$

- *A 는 촬영된 영상*
- ** 는 2D convolution 연산자*

2.3 TSV Hole 영상 추출

TSV Hole이 촬영된 영상에는 Via Hole외에 칩을 구성하고 있는 다양한 구조물도 포함되어 있다. 본 논문에서는 Via Hole 영역만 잘라내기 위해 탬플릿 매칭(template matching) 기법을 사용하였다. 임의의 Via Hole 영역을 42x42 크기로 잘라낸 탬플릿 영상과 촬영 영상내의 모든 부분 영상(sub image)들 간 NCC(normalized cross correlation)를 계산하고(식 (2) 참조)[7], NMS(non-maximum suppression)을 적용해Via Hole이 포함된 부분 영상을 추출하였다.

$$NCC = \frac{1}{n}\sum_{x,y}\frac{(f(x,y)-\bar{f})(t(x,y)-\bar{t})}{\sigma_f \sigma_t} \quad (2)$$

- *n 은 화소의 개수*
- *f(x,y)는 부분 영상의 [x,y] 화소*
- *t(x,y)는 탬플릿 영상의 [x,y] 화소*

2.4 Ground Truth 설정

각 TSV Hole에 대해 CD-SEM 장비의 지름 측정 기능을 적용하여 계측한 결과를Ground Truth로 사용하였다.

2.5 TCD 측정 딥러닝 모델

딥러닝 모델 중 CNN은 영상 인식, 분류 등에 많이 사용되는 모델이다. CNN의 레이어는 크게 Conv(convolution)/Pooling 레이어와 FC(fully connected) 레이어로 나뉜다. Conv 레이어는 영상에서 특징점을 추출한다. FC 레이어는 Conv 레이어에서 추출된 특징점을 기반으로 영상을 분류한다.

통상적인 기존 연구에서는 높은 성능을 내는 기존 CNN모델을 주어진 과제에 맞게 수정하여 적용하고 성능을 평가하는 식으로 진행된다[9-11]. 그러나, 본 논문에서 해결하고자 하는 문제는 42x42 크기의 입력 영상에 대한 회귀(regression) 예측이기 때문에 보다 단순한 모델로도 해결이 가능하다고 판단했다.

본 논문에서는 Via Hole의 TCD 측정을 위한 최적의 CNN 모델을 도출하기 위해 Conv 레이어를 점진적으로 추가했을 때의 성능 변화와 FC레이어를 추가하였을 때의 성능 변화를 관찰하였다. Weight의 초기값과 데이터 세트의 변화로 인한 성능 변화를 막기 위해 Random함수의 Seed 값을 임의의 수로 고정하고 학습과 테스트를 진행하였다. 모든 평가에서 학습률(learning rate)은 0.001, batch size는 1로 설정하였다. 최적화 방법으로 Adam optimizer[12]를 사용하였다. L2 규제(regularization)의 강도는 0.00001을 적용하였고, 조기 종료(early stopping) 기법을 적용하여 과적합(overfitting)을 방지하였다. 모델의 손실 함수로 평균 제곱오차(mean squared error)를 사용하였다. 모든 레이어 간의 활성함수는 Relu를 사용하였다. 입력이 Conv레이어를 통과할 때마다Max Pooling을 적용하여 크기를 ¼씩 줄여나갔다.

본 논문에서의 CNN 모델 탐색 과정은 다음과 같다.

1) Conv 레이어 구성 탐색

Conv레이어의 수와 각 Conv레이어의 크기에 대한 탐색 과정은 다음과 같다.

a) Conv레이어 크기를 1x1x1, Stride=1, Padding=0, Bias=True 로 설정
b) 학습과 테스트 진행
c) Depth 1 증가 후 학습과 테스트 진행
d) Depth가 21이 될 때까지 단계 c 반복
e) Filter 크기 2 증가, Depth 1로 설정
f) Filter 크기가 Convolution 연산이 불가능할 때까지 단계 b~e까지 반복
g) 단계 a~f에서 최고의 성능을 갖는 Conv레이어 선택
h) 단계 g에서 선택된 Conv레이어 뒤에 신규Conv레이어 추가
i) 신규 Conv레이어에 대해 단계 a 부터 h까지 반복

본 논문에서는 Conv레이어 4개 구성까지 탐색하였고 그 결과 Conv1 레이어 7x7x3, Conv2 레이어 7x7x13, Conv3 레이어 5x5x9로 구성될 때의 성능이 가장 높게 나오는 것을 확인하였다.

2) FC 레이어 구성 탐색

FC레이어 수에 대한 탐색 과정은 다음과 같다.

a) 첫 번째 FC레이어의 입력은 9로 설정
b) 학습과 테스트 진행
c) 이전 FC레이어의 ½ 크기인 FC레이어를 추가 후 학습과 테스트 진행
d) 추가되는 FC레이어의 크기가 1보다 작을 때까지 단계 c 반복

본 논문에서는 FC레이어 1개를 사용한 경우의 성능이 가장 좋았고, FC레이어가 더 추가되어도 성능이 좋아지지 않았다.

3) CNN 모델 구성

단계 1)과 2)를 통해 도출된 CNN 모델의 구조는 다음과 같다.

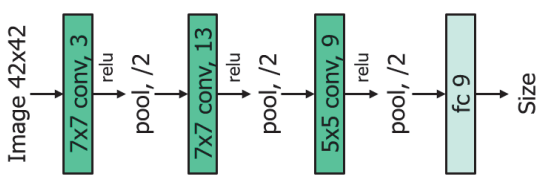

Fig. 5. Proposed Deep Learning Model.

3. 실험 및 고찰

3.1 데이터 세트

실험을 위해 임의의 Via Hole 영상 2461개를 데이터 세트로 사용하였고, CD-SEM으로 측정한 해당 데이터 세트의 TCD 분포는 Fig. 6과 같다. 실험에 사용된 시료의TCD 정상 범위는 5.6 ± 0.3 um이다. 학습/검증(validation)/테스트 데이터 세트 비율은 0.8:0.1:0.1이다.

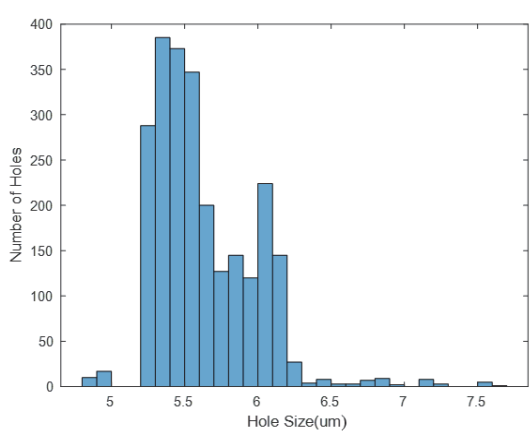

Fig. 6. Distribution by Size of Data Set.

3.2 실험

본 논문에서 제안한 모델의 효용성을 입증하기 위해 ResNet18 모델의 결과와 비교하였다. ResNet18 모델의 입력 크기인 224 x 224를 맞추기 위해 데이터 세트 크기를 42 x 42에서 224 x 224로 Resize하였고, Resize시에는 bilinear interpolation 기법을 사용하였다. ResNet18의 학습률을 0.001로 했을 때, 그레디언트 폭주(Gradient exploding) 현상이 발생하여 0.00001로 줄여서 학습을 진행하였다. 학습되는 파라미터 수는 ResNet18이 약 1151만개, 제안한 모델은 5018개이다.

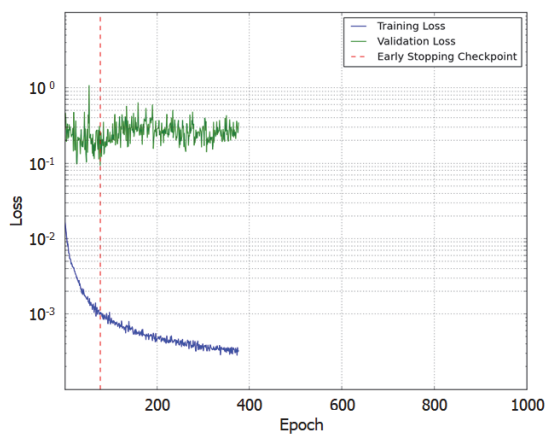

Fig. 7. ResNet18 Training & Validation.

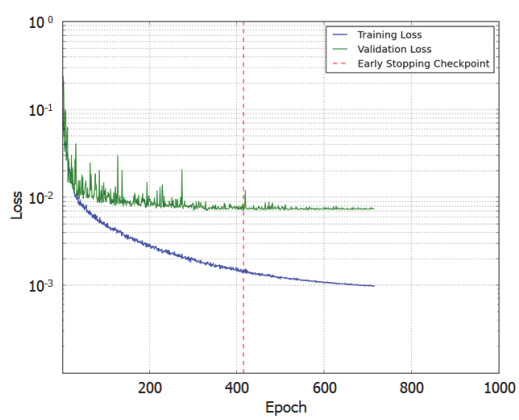

Fig. 8. Proposed Model Training & Validation.

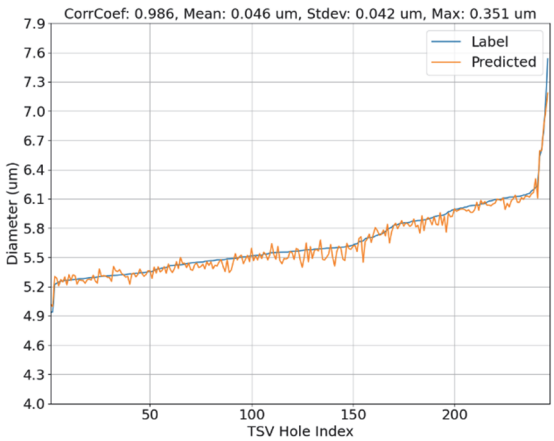

Fig. 9. ResNet18 Prediction Result.

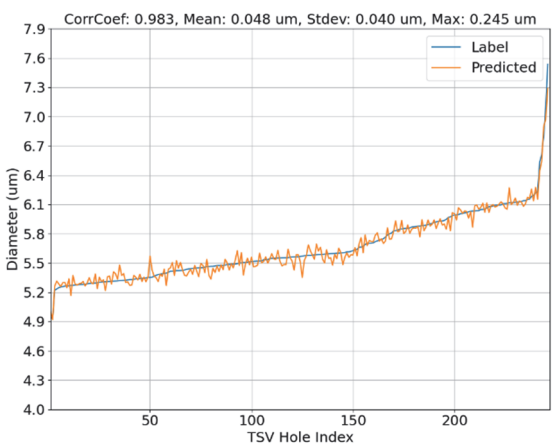

Fig. 10. Proposed Model Prediction Result.

Fig. 7과 Fig. 8은 각각 ResNet18과 제안 모델의 학습 결과이다. 두 모델 모두 Early Stopping이 적용되어 500 epoch 이전에 학습이 끝났다.

Fig. 9와 Fig. 10은 테스트 데이터 세트에 대해 각 모델이 TCD를 예측한 결과이다. 실험 결과 두 모델의 예측값 모두 CD-SEM 측정값과 약 98%이상의 상관도를 보였으나 ResNet18이 0.3% 높았다. 평균(mean)과 표준 편차(stdev) 그리고 최대값(max)은 CD-SEM 측정값과 모델의 예측값 간 절대차(absolute difference)를 기반으로 계산되었다. 평균은 ResNet18이 0.002 um 작았으나 표준 편차와 최대값은 제안한 모델의 예측 결과가 더 작았다.

Table 1은 제안한 모델의 재현성(reproducibility)을 검증하기 위해 Random 함수의 Seed 값을 고정하지 않고 학습과 테스트를 6회 추가적으로 진행해 얻은 결과이다. 상관도 기준으로 최대 최소 결과는 제외하였다.

Table 1. Reproducibility of proposed model results

	CorrCoef (%)	Mean (um)	Stdev (um)	Max Diff (um)
1	0.981	0.048	0.046	0.381
2	0.979	0.057	0.046	0.233
3	0.976	0.052	0.048	0.293
4	0.982	0.048	0.039	0.244
5	0.982	0.051	0.048	0.342
Avg.	0.980	0.051	0.045	0.299

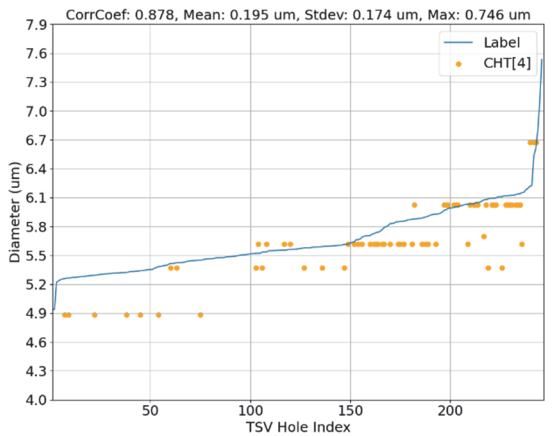

Fig. 11. Circle Hough Transform Result.

Table 1을 살펴보면 추가 실험의 평균 결과는 Fig. 10의 결과와 크게 차이 나지 않았다. 따라서 제안한 모델의 성능이 재현성을 보임을 확인할 수 있다.

Fig. 11은 테스트 데이터 세트에 대해 [4]를 적용한 결과이다. 246개의 데이터 중 67개에서 원을 검출할 수 있었다. 알고리즘의 파라미터가 최적화되지 않았다는 점을 감안해도 제안한 모델의 성능보다 크게 못 미치는 것을 확인할 수 있다.

4. 결 론

본 논문에서는 TSV Hole의 TCD 측정을 위한 딥러닝 모델을 제안하였고, 기존 딥러닝 모델인 ResNet18과 성능을 비교하였다. 제안된 모델의 학습 파라미터 수는 기존 모델보다 약 1/2000 적음에도 계측 결과는 기존 모델의 결과와 대등했다. 실험 결과CD-SEM 계측값과 제안한 모델의 예측값의 상관도가 평균적으로 98%, 평균 차이는 0.051 um, 표준 편차는 0.045 um, 최대 차이는 0.299 um로 기존 알고리즘 방식보다 성능이 크게 개선되었음을 확인하였다.

참고문헌

1. Victor H. Vartanian, Richard A. Allen, Larry Smith, Klaus Hummler, Steve Olson, and Brian C. Sapp, “Metrology needs for through-silicon via fabrication”, Journal of Micro/Nanolithography, MEMS, and MOEMS 13(1), 011206 (27 February 2014).
2. Arun A. Aiyer, Tianheng Wang, and Helen Simson, “Optical inspection technologies for 3D packaging”, https://www.researchgate.net/publication/283489453_Optical_Inspection_Technologies_for_3D_Packaging, 2015.
3. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Deep Residual Learning for Image Recognition”, arXiv:1512.03385v1 [cs.CV], 2015.
4. T.J Atherton, D.J. Kerbyson, “Size invariant circle detection”, Image and Vision Computing. Volume 17, Number 11, 1999, pp. 795-803.
5. Ashardi Abas, Xun Chen, “COIL CHAMFER DIAMETER MEASUREMENT IN HELICAL SPRING BASED ON A HUGH TRANSFORM ALGORITHM”, 4th International Conference on Artificial Intelligence and Computer Science, November 2016.
6. Ali Ajdari Rad, Karim Faez, and Navid Qaragozlou, “Fast Circle Detection Using Gradient Pair Vectors”, Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney.
7. J. P. Lewis, “Fast Template Matching”, Vision Interface, pp. 120-123, 1995.
8. Dirk-Jan Kroon, “Numerical Optimization of Kernel Based Image Derivatives”, Short Paper, University of Twente, 2009.
9. Sung Joo Kim, Kim Gyung Bum, “A Study on the Classification of Surface Defect Based on Deep Convolution Network and Transfer-learning”, Journal of the Semiconductor & Display Technology, Vol. 20, No. 1, March 2021.
10. Sung-jin Hwang, Seok-woo Hong, Jong-seo Yoon, Heemin Park, Hyun-chul Kim, “Deep Learning-based Pothole Detection System”, Journal of the Semiconductor & Display Technology, Vol. 20, No. 1, March 2021.
11. Song-Yeon Lee, Yong Jeong Huh, “A Study on Shape Warpage Defect Detecion Model of Scaffold Using Deep Learning Based CNN”, Journal of the Semiconductor & Display Technology, Vol. 20, No. 1, March 2021.
12. Diederik P. Kingma, Jimmy Ba, “Adam: A Method for Stochastic Optimization”, Proceedings of the 3rd International Conference on Learning Representations, 2015.

접수일: 2021년 6월 10일, 심사일: 2021년 6월 17일, 게재확정일: 2021년 6월 22일