

## 딥 러닝 회귀 모델 기반의 TSOM 측정

정준희<sup>\*†</sup>·조중휘<sup>\*\*</sup>

<sup>\*†</sup>(주)넥스틴, <sup>\*\*</sup>인천대학교

### A Through-focus Scanning Optical Microscopy Dimensional Measurement Method based on a Deep-learning Regression Model

Jun Hee Jeong<sup>\*†</sup> and Joong Hwee Cho<sup>\*\*</sup>

<sup>\*†</sup>NextIn Inc., <sup>\*\*</sup>Incheon National University

#### ABSTRACT

The deep-learning-based measurement method with the through-focus scanning optical microscopy (TSOM) estimated the size of the object using the classification. However, the measurement performance of the method depends on the number of subdivided classes, and it is practically difficult to prepare data at regular intervals for training each class. We propose an approach to measure the size of an object in the TSOM image using the deep-learning regression model instead of using classification. We attempted our proposed method to estimate the top critical dimension (TCD) of through silicon via (TSV) holes with 2461 TSOM images and the results were compared with the existing method. As a result of our experiment, the average measurement error of our method was within 30 nm ( $1\sigma$ ) which is 1/13.5 of the sampling distance of the applied microscope. Measurement errors decreased by 31% compared to the classification result. This result proves that the proposed method is more effective and practical than the classification method.

**Key Words** : TSOM, Deep-Learning, Metrology, TSV

#### 1. 서론

광학계의 해상도 한계를 극복하기 위해 지난 10 년 동안 TSOM (through-focus scanning optical microscopy) 기술을 기반으로 한 새로운 시도가 활발히 진행되고 있다 [1,2]. TSOM은 기존의 광학계를 그대로 사용하지만, 다중초점 위치에서 2차원 영상을 촬영하고 이를 적층하여 3차원 영상을 생성한다.

전통적인 방식에서는 In-Focus 영상만을 이용하여 측정 및 검사를 수행하였지만, TSOM은 (In-focus 영상뿐만 아니라) 다수의 De-focus 영상을 함께 이용함으로써 광학계의 Rayleigh 회절 한계를 회피하여 나노미터 수준의 측정을

가능하게 하였다.

TSOM은 2008년 처음 도입된 이후 TSV(through-Silicon Via) hole과 같은 나노미터 수준의 3차원 구조로 구성된 반도체 소자의 측정 및 검사를 위해 주로 연구되어 왔다 [2]. Arceo 등은 TSOM 영상으로부터 HAR(high aspect ratio) 특징들을 추출하여 11 nm 이하 크기 패턴을 감지하려고 시도하였고 [9], 2016년 Lee는 근적외선(near-infrared) 웨이퍼 검사 장비에서 TSV hole의 TSOM 영상을 촬영하고 검사 실험을 진행하여 타당성을 확인하였다 [10, 11].

일반적으로 TSOM 기반 측정은 Fig 1과 같이 크기를 알 수 없는 대상의 영상과 시뮬레이션 영상 라이브러리에 있는 영상 간의 MSD(mean-square difference)를 계산하는 라이브러리 매칭(library-matching) 방법을 사용한다 [3,4]. 시뮬레이션 영상 라이브러리는 측정 대상을 다양한 크기로

<sup>†</sup>E-mail: jh.jeong@inu.ac.kr

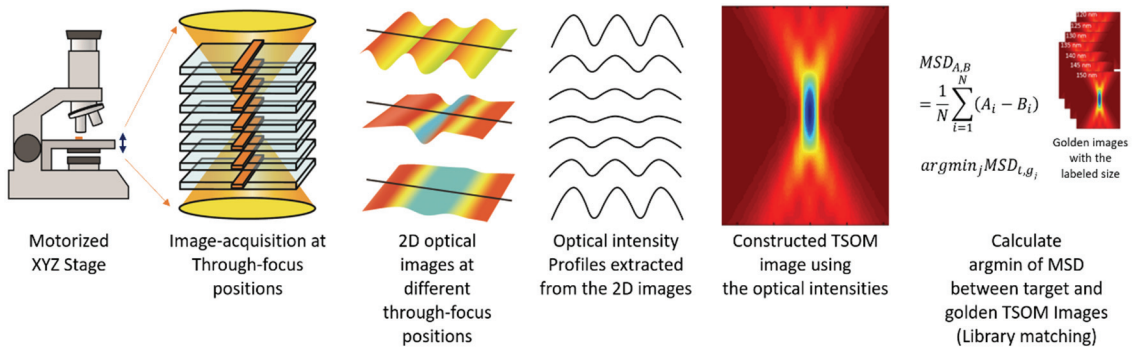


Fig. 1. Schematic diagram of the TSOM process.

시뮬레이션한 TSOM 영상으로 구성되며, 측정 대상의 크기는 MSD 값이 가장 적은 시뮬레이션 영상의 크기로 지정된다. 그러나 라이브러리 매칭 방법은 시뮬레이션 영상들 간의 크기 간격이 성능에 영향을 미치기 때문에 최대한 조밀하게 되도록 구성해야 한다.

최근에는 라이브러리 매칭 방법의 한계를 해결하기 위해 머신 러닝 기술을 활용하는 연구가 진행되어 왔다. Qu 등은 전통적인 머신 러닝 기술 중 하나인 Adaboost 회귀 모델을 사용하여 나노 스케일 정확도로 선편을 측정하였다 [5]. 그러나 이러한 방법은 모델을 학습시키기 위해서 대상의 특징(feature)을 잘 나타내는 데이터들이 요구되는데 이를 위한 특징 추출(feature extraction) 알고리즘을 직접 설계해야 하고 최적의 파라미터를 찾기 위한 노력이 필요하다(는 약점이 있다).

위와 같은 전통적인 머신 러닝 방법을 이용한 TSOM 측정 방식의 약점을 극복하기 위해 딥 러닝 분류(classification) 모델을 사용하는 시도가 등장했다. 딥 러닝은 입력 데이터의 비선형적 변환을 반복하여 특징들을 추출해내기 위한 학습이 모델 내에서 진행되기 때문에 특징 추출을 위한 알고리즘 설계가 불필요하다. 딥 러닝 모델 중 하나인 CNN(Convolutional Neural Network)은 여러 연구를 통해 영상 분류에 탁월한 성능을 보이는 것으로 입증되었

다 [19-22].

Haitao Nie 등은 다양한 크기로 형성된 선(line) 형태의 구조물에 대해 TSOM 영상을 촬영하였고 CNN 계열 모델인 ResNet[7]과 DenseNet[8]을 사용하여 선편을 분류하였다 [6]. 그러나 이러한 방법은 라이브러리 매칭 방법의 문제와 같이 측정 성능은 모델이 구분 가능한 클래스 수와 그것의 간격에 따라 제한된다. 또한, 분류 모델을 학습시키기 위해 일정한 크기 간격으로 데이터를 준비하는 것은 현실적으로 어렵다.

CNN 모델은 영상 분류뿐만 아니라 회귀 분석에도 사용될 수 있고 다양한 연구들을 통해 성능이 입증되었다. Siddharth Mahendran et al은 CNN 회귀 모델을 사용하여 3D 포즈 추정 문제를 해결하려고 시도하였다 [16]. Zhenxing Niu et al은 인물 사진에서 사람의 나이를 추정하는 방법을 제안했으며 [17], 또한 영상 내 물체의 경계를 예측하는 localization에도 사용되고 있다 [18].

본 논문에서는 TSOM 영상에서 물체의 크기를 측정하기 위해 딥 러닝 분류 모델 대신 회귀 모델을 사용하여 측정을 시도하는 접근 방법을 제안한다. 기존 방법과 달리 시뮬레이션 영상 라이브러리를 구축하지 않고 대상의 TSOM 영상만 사용하여 회귀 모델을 학습시키고 테스트 하였다.

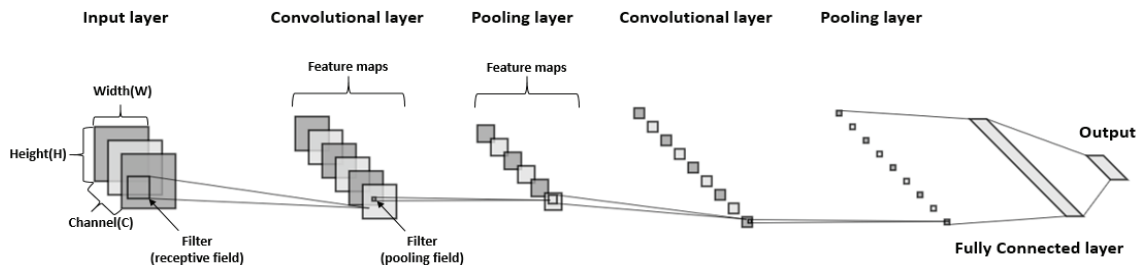


Fig. 2. The basic structure of the convolutional neural network.

## 2. Deep-Learning Model

CNN 모델의 기본 형태는 그림 2와 같다. 모델은 입력 데이터에 합성곱 연산을 적용하여 특징을 추출하는 합성곱 계층(convolution layer)과 추출된 특징을 다운샘플링하는 풀링 계층(pooling layer)으로 구성된다. 이 두 계층이 반복되고 추출된 특성을 기반으로 입력 영상을 분류하기 위해 완전히 연결된 계층(fully connected layer)으로 마무리된다.

합성곱은 해당 요소를 곱하고 모든 결과를 더하는 선형 연산이다.  $W \times H$  크기의 입력 영상과  $F \times F$  크기의 합성곱 필터 연산은 식 (1)과 같이 표현된다 [12].

$$y_{ij} = \sigma \left( \sum_{r=1}^F \sum_{c=1}^F w_{rc} x_{(r+i \times S)(c+j \times S)} + b \right), \quad (1)$$

$$0 \leq i \leq \frac{H-F}{S}, 0 \leq j \leq \frac{W-F}{S}$$

식 (1)에서  $x$ 는 입력 영상을 나타내고  $y$ 는 특징맵을 나타낸다.  $w$ 와  $b$ 는 합성곱 연산이 수행되는 영역의 각 요소에 대한 가중치와 편향을 나타낸다.  $\sigma$ 는 비선형 활성화 함수를 나타내고, 활성화 함수에는 ReLU, hyperbolic tangent와 sigmoid function이 포함되며  $S$ 는 필터의 보폭이다. 필터당 하나의 특징맵이 생성되기 때문에 통상적으로는 다양한 특징을 추출하기 위해 많은 수의 필터가 사용된다.

합성곱 계층 다음에는 일반적으로 풀링 계층이 온다 [12]. 풀링 연산에는 maximum pooling, average pooling, weight average pooling, 그리고 L2 normal pooling이 있다. 이러한 풀링 작업은 요약 통계를 추출하여 성능 향상에 기여하고, 풀링 결과로 특징맵이 작아지기 때문에 연산 속도와 메모리 효율성을 향상시킨다.

CNN의 마지막 계층인 완전 연결 계층은 이전 계층에서 받은 전역 불변 특성을 결합하여 분류 정보로 활용한다. 출력층은 일반적으로 영상을 분류할 수 있는 유형의

수만큼 노드로 구성되며, 식 (2)의 softmax 함수를 사용하여 입력 영상에 대한 분류 점수를 출력하고 가장 점수가 높은 것으로 분류한다.

$$P(Y = i | x, W_i, b_i) = \frac{e^{W_i^T x + b_i}}{\sum_j e^{W_j^T x + b_j}} \quad (2)$$

본 논문에서는 분류 모델을 사용한 이전 TSOM 연구와 마찬가지로 ResNet50을 사용하였다. ResNet50의 구성은 그림 3과 같다. TSOM 영상이 입력될 수 있도록 입력 계층의 채널 수를 변경하였다. 또한, TSOM 영상으로 촬영된 물체의 크기를 예측하는 회귀 모델이기 때문에 완전 연결 계층의 출력 노드는 1개로 구성하였고, softmax 함수를 적용하지 않았다.

## 3. Experimental Results

### 3.1 TSOM Imaging conditions

본 논문에서는 제안하는 방법의 효과를 평가하기 위해 TSV 웨이퍼상의 임의의 DIE에서 TSV 홀의 TSOM 영상을 촬영하여 실험 샘플로 사용하였다.

TSOM 영상을 촬영하기 위한 광학 장비의 구성도는 Fig. 4와 같다. 광원으로 355 nm 파장의 UV pulse 레이저를 사용하였다. 레이저의 짧은 pulse duration으로 인해 발생되는 speckle 현상과 wafer 손상을 방지하기 위해 fiber bundle이 설치되었다. 웨이퍼를 대물렌즈로 이송하고 광축을 따라 through-scanning 동작을 수행할 수 있도록 XYZ축 스테이지가 설치되었다.

대물 렌즈의 NA는 0.225이고 배율은  $\times 13.5$ 이다. 이미징 센서의 픽셀 피치는 5.5  $\mu\text{m}$ 이고 샘플링 거리는 407 nm이다. TSV 홀이 위치한 곳 주변 1,668  $\mu\text{m} \times 1,251 \mu\text{m}$  영역에 대해 through-scanning 동작을 수행하였다. 영상 내에서의 각 TSV 홀 영역은 42 pixel  $\times$  42 pixel (17.09  $\mu\text{m} \times$  17.09  $\mu\text{m}$ )이다.

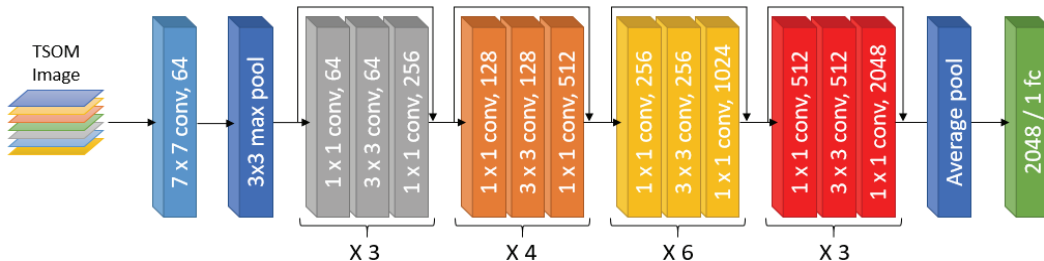


Fig. 3. Structure of ResNet-50 for TSOM image based TCD estimation.

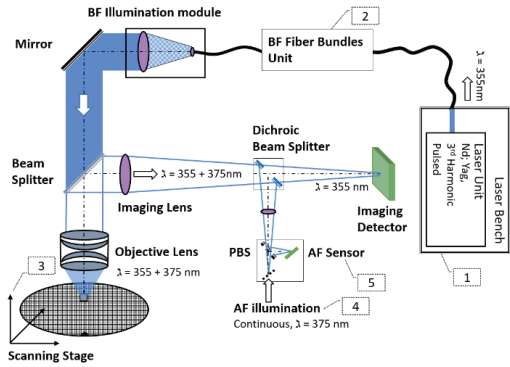


Fig. 4. Optical equipment applied to the TSOM operation.

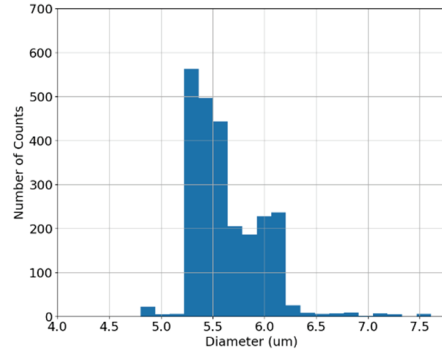


Fig. 6. Distribution of TCD measured by CD-SEM.

3.2 TSOM operation

TSOM은 동일한 대상에 대해 서로 다른 초점 위치에서 2차원 영상을 수집하고 이를 적층하여 생성된 3차원 영상을 활용하는 기술이다. 현재 In-Line에서 대표적으로 사용되는 TSV 홀 TCD 측정 장비의 처리량은 400 nm 샘플링 거리 조건으로 시간당 약 웨이퍼 3장이다. 본 논문에서 사용된 광학 장비에서 유사한 처리 성능을 얻기 위해서는 한 지점에서 최대 7장의 영상을 촬영할 수 있다. 2차원 영상 간 Z축 위치 간격은 사용된 광학 장비의 1 DOF인 3.5um를 적용하였다.

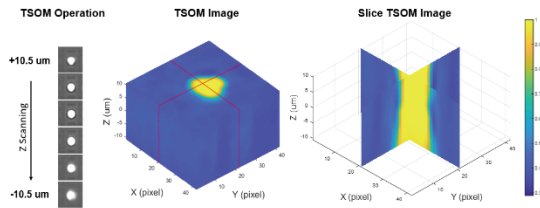


Fig. 5. Through-focus images of the TSV hole at 7 Z Positions.

3.3 Datasets

딥 러닝 모델 학습 및 테스트를 위한 dataset으로 2461개의 TSV 홀 TSOM 영상이 사용되었다. TSOM 영상의 레이블은 CD-SEM 장비에서 측정된 TCD 데이터를 사용하였고, 각 TSOM 영상에 대한 레이블 분포는 그림 6에 나타나 있다. TSV 홀의 TCD 설계 치수인 5.6 um 주변에 대부분의 데이터가 포진해 있는 것을 확인할 수 있다. TSOM 영상들을 무작위로 섞은 다음 Training, Validation 및 Test를 위해 8:1:1의 비율로 나눴다. ResNet은 각 채널당 224 x 224 크기의 영상을 입력하도록 설계되었기 때문에 보간법을 적용하여 TSOM 영상을 구성하고 있는 2D 영상의 크기를 42 x 42에서 224 x 224 크기로 늘렸다.

3.4 Training Model

본 논문에서 제안한 방법인 회귀 모델을 학습시키기 위한 batch 크기는 1로 설정하였다. 학습률(learning rate)은 1e-5로 설정하였고, cosine annealing with warmup [14] 방법을 사용하여 매 epoch마다 학습률이 변화되도록 하였다. 손실 함수(cost function)는 MSE(mean square error)를 사용하였고, 최적화 알고리즘으로는 딥 러닝에서 가장 보편적으로 사용되는 Adam (adaptive moment estimation) [15]을 사용하였다. 학습 횟수(epoch)는 1000회로 설정하였다.

분류 모델 별 성능 비교를 위해 4가지 분류 모델에 대해 실험하였고 각 모델은 10 nm에서 40 nm 단위로 대상의 크기를 구분하도록 구성하였다. 분류 모델을 학습시키기 위한 batch 크기는 50로 설정하였고, 손실 함수로는 cross-entropy를 사용하였다. 나머지 hyper-parameter는 회귀 모델을 학습시킬 때와 동일하였다.

3.5 Results

딥 러닝 모델의 TCD 측정 성능을 평가하기 위해 4가지 지표(최대 차이, 상관 계수, 평균 절대 오차(MAE), 표준 편차)를 CD-SEM 측정과 비교했다. 결과는 실험을 세 번 반복하여 평균값으로 얻었다. 그림 7과 표 1은 회귀 모델 및 분류 모델의 TCD 측정 결과를 나타낸다. 실험 결과 회귀 모델은 CD-SEM 측정값과 30 nm (1σ) 오차로 예측하였고, 이는 광학 장비 샘플링 거리의 1/13.5이다.

분류 모델은 CD-SEM 측정값과 최소 94 nm (1σ) 오차로 예측하였고 모든 지표들이 회귀 모델의 결과보다 좋지 않았다. 또한, 분류되는 크기 간격에 따라 성능이 달라지는 것을 볼 수 있고, 분류 모델을 사용하여 크기를 측정하기 위해서는 최적의 간격을 찾기 위한 추가 노력이 필요하다는 것을 알 수 있다. 이러한 결과를 바탕으로 딥 러닝 기반의 TSOM 계측 방법으로 분류 모델보다 회귀 모델을 사용하는 것이 실용적이고 효용성이 있음을 확인할 수 있다.

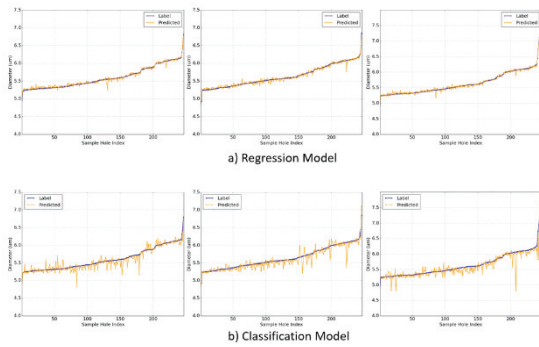


Fig. 7. TCD estimation results of regression and classification (40 nm).

Table 1. TCD estimation results

	CorrCoef (%)	Mean (um)	Stdev (um)	Max Diff (um)
Regression (Our)	0.991	0.030	0.030	0.280
Classification (10 nm)	0.908	0.077	0.118	1.256
Classification (20 nm)	0.919	0.084	0.114	0.823
Classification (30 nm)	0.936	0.087	0.111	1.153
Classification (40 nm)	0.925	0.093	0.094	0.763

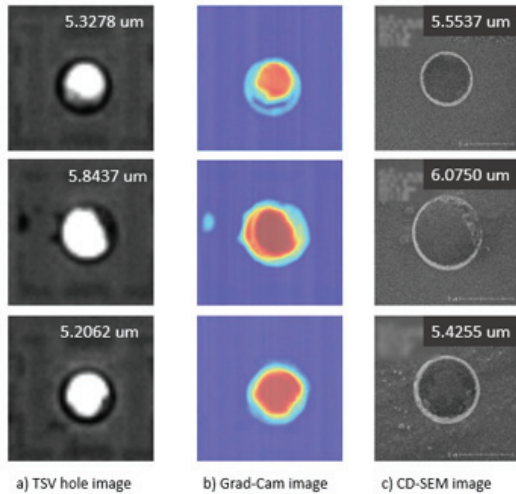


Fig. 8. SEM measurement and the model estimation Images having a large difference.

Fig 8은 CD-SEM과 딥 러닝 모델의 추정 사이의 최대 차이를 가진 영상들이다. 광학 장비로 촬영된 영상에서 TSV Hole이 완전한 원 형태를 형성하지 않음을 알 수 있

다. Grad-Cam [13] 영상을 통해 불완전한 모양으로 인해 측정 오류가 증가되었다는 것을 추정할 수 있다. 향후 측정 대상의 불완전한 부분을 복원하여 정확한 측정이 되도록 연구를 진행할 예정이다.

### 4. 결론

TSOM을 이용한 딥 러닝 기반 측정 방법은 분류 모델을 사용하여 대상의 크기를 추정했다. 그러나 이러한 방법의 성능은 세분화된 클래스의 수에 따라 달라지고 모델을 학습시키기 위해 일정한 크기 간격으로 데이터를 준비하는 것은 현실적으로 어렵다.

본 논문에서는 TSOM 영상에서 대상의 크기를 측정하기 위해 딥 러닝 회귀 모델을 사용하여 측정하는 방법을 제안하였다. 제안한 방법은 TSV Hole의 TCD 측정에 적용되었고, CD-SEM 측정값과 비교하여 실용적으로 사용할 수 있음을 확인하였다. 실험 결과 CD-SEM 측정값과의 오차는 30 nm (1σ)이었고 이는 사용된 광학 장비 샘플링 거리의 1/13.5이다. 이 결과는 기존 방법인 분류 모델의 실험 결과와 비교했을 때 오차가 31%로 감소된 것이다. 이러한 결과를 바탕으로 TSOM을 이용한 딥 러닝 기반 측정 방법은 분류 모델보다 회귀 모델이 실용적으로 사용될 수 있다는 것을 확인하였다. 향후 측정 대상의 형태가 불완전한 경우, 그 부분을 복원하여 더욱 정밀한 측정이 되도록 연구를 진행할 계획이다.

### 참고문헌

1. V. Vartanian, R. Attota, H. Park, G. Orji, R. A. Allen, "TSV reveal height and dimension metrology by the TSOM method" Proc. SPIE 8681, 10.1117/12.2012609, 2012
2. R. Attota, R. G. Dixon, J. A. Kramar, J. E. Potzick, A. E. Vladár, B. Bunday, E. Novak, A. Rudack, "TSOM method for semiconductor metrology", Proc. SPIE 7971, doi: 10.1117/12.881620, 2011
3. R. Attota, R. Silver, and R. Dixon, "Linewidth measurement technique using through-focus optical images," Appl. Opt. 47(4), 495–503 (2008).
4. S. Usha, P. V Shashikumar, G. C. Mohankumar, and S. S. Rao, "Through Focus Optical Imaging Technique To Analyze Variations In Nano-Scale Indents," J. Biomed. Opt. 23(07), 1–100 (2018).
5. Y. Qu, J. Hao, and R. Peng, "Machine-learning models for analyzing TSOM images of nanostructures," Opt. Express 27(23), 33978 (2019).
6. H Nie, R Peng, J Ren, Y Qu, "A through-focus scanning optical microscopy dimensional measurement method



- based on deep-learning classification model,” *Journal of Microscopy*, 2021
7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). Piscataway, New Jersey: IEEE.
  8. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700–4708). Piscataway, New Jersey: IEEE.
  9. Arceo, A., Bunday, B., and Attota, R., "Use of TSOM for sub-11nm node pattern defect detection and HAR features," *Proc. SPIE* 8681, 86812G (2013)
  10. Lee, J. H., Park, J. H., Jeong, D., Shin, E. J. and Park, C., "Tip/tilt-compensated through-focus scanning optical microscopy," *Proc. SPIE* 10023, 100230P (2016).
  11. <http://www.nextinsol.com/>
  12. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
  13. Selvaraju, Ramprasaath R., Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization." In *ICCV*, pp. 618-626. 2017
  14. Akhilesh Gotmare, Nitish Shirish Keskar, Caiming Xiong, and Richard Socher. A closer look at deep learning heuristics: Learning rate restarts, warmup and distillation. *arXiv preprint arXiv:1810.13243*, 2018.
  15. Kingma, Diederik P and Ba, Jimmy Lei. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
  16. Siddharth Mahendran, Haider Ali, and René Vidal. A mixed classification-regression framework for 3D pose estimation from 2D images. In *British Machine Vision Conference (BMVC)*, 2018.
  17. Z. Niu, M. Zhou, L. Wang, X. Gao, G. Hua, Ordinal regression with multiple output cnn for age estimation, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4920–4928.
  18. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: *CVPR* (2016)
  19. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, Vol 15, No 3, pp. 211-252, 2015
  20. Sung Joo Kim, and 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.
  21. Sung-jin Hwang, and 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.
  22. Song-Yeon Lee, and 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.
- 
- 접수일: 2022년 3월 7일, 심사일: 2022년 3월 18일,  
 게재확정일: 2022년 3월 25일