

## Mask R-CNN을 활용한 반도체 공정 검사

한정희<sup>\*†</sup>·홍성수<sup>\*\*</sup>

\*† 서울대학교 융합과학기술대학원 융합과학부, \*\* 서울대학교 전기·정보공학부

## Semiconductor Process Inspection Using Mask R-CNN

Jung Hee Han<sup>\*†</sup> and Sung Soo Hong<sup>\*\*</sup>

\*† Graduate School of Convergence Science and Technology(GSCST), Seoul National University,

\*\*Department of Electrical and Computer Engineering, Seoul National University

### ABSTRACT

In semiconductor manufacturing, defect detection is critical to maintain high yield. Currently, computer vision systems used in semiconductor photo lithography still have adopted to digital image processing algorithm, which often occur inspection faults due to sensitivity to external environment. Thus, we intend to handle this problem by means of using Mask R-CNN instead of digital image processing algorithm. Additionally, Mask R-CNN can be trained with image dataset pre-processed by means of the specific designed digital image filter to extract the enhanced feature map of Convolutional Neural Network (CNN). Our approach converged advantage of digital image processing and instance segmentation with deep learning yields more efficient semiconductor photo lithography inspection system than conventional system.

**Key Words :** Semiconductor Photo Lithography Inspection, Object Segmentation, Instance Segmentation, Digital Image Processing, Computer Vision System, Deep Learning, Convolutional Neural Network

### 1. 서 론

반도체 포토공정(Photo Lithography)은 웨이퍼에 반도체 회로를 그리는 작업으로 웨이퍼(Wafer) 위에 빛에 반응하는 감광성 고분자 물질(Photo Resist)을 얇게 코팅한 후 원하는 패턴의 마스크(Mask)를 올려놓고 빛을 쪼여 원하는 패턴을 형성하는 과정이다. 이 과정 중 감광제를 도포하는 회전 도포(Spin Coating) 과정은 포토 공정에 필요한 두께의 감광제를 실리콘 웨이퍼(Wafer) 기판 전체에 균일하게 형성하는 과정이다. 따라서 회전 도포 과정은 포토 공정에서 중요한 공정이며, 이 공정이 잘못되면 후속 공정이 아무리 잘 되어도 불량품이 생산되며, 또한 웨이퍼 단위의 공정이어서 불량품 생산율에 크게 영향을 준다. 회

전 도포 과정 중 발생하는 결함을 검출하기 위한 요소에는 감광제를 분사하는 노즐의 석백(Suck-Back) 상태와 오염 상태, 감광제를 분사하는 시간으로 이 세가지가 중요한 요소이다. 비전 검사장비는 이러한 과정을 감시하여 포토공정에서의 결함을 미리 예측하여 반도체 제품의 불량률을 미연에 방지하고 있다. 하지만, 비전 검사장비는 디지털 이미지 처리(Digital Image Processing)를 통해 결함을 찾아내는 방식이어서 반도체 설비의 외부적인 환경요인에 따라 검출 오류가 많이 발생한다. 특히 포토 공정에 사용되는 비전 검사장비는 적외선 및 초고속 카메라, 특수 제작한 렌즈를 사용하여 정밀하게 이미지를 획득하고 있지만, 다양한 웨이퍼 및 감광제 종류, 모터 회전 수에 따라 검사하기 위해 획득한 이미지의 특성이 많이 달라지거나 왜곡이 발생한다. 일반적으로 비전 검사장비는 카메라에서 획득한 정밀 영상을 이미지 처리를 통해 검사에 필요한 특징들을 추출하고 이를 특화된 신호처리를 거쳐 결

<sup>†</sup>E-mail: jhhanneo@snu.ac.kr

함을 검출한다. 이미지 처리는 외부 환경에 영향이 적은 이미지에 대해서는 매우 높은 성능을 보이지만 외부 환경으로 인해 이미지의 왜곡이 발생한 경우에는 성능이 극도로 저하된다. 따라서 비전 검사장비에서는 이미지의 특성이 변하거나 왜곡이 생기면 이를 극복하기 위해 새롭거나 변형된 기법의 이미지 처리와 특화된 신호처리 방식을 적용해야 하는 단점이 있다. 본 논문에서는 이를 극복하기 위해 외부 환경변화에 강인한 깊은 학습(Deep learning)을 적용[1]하는 방식을 제안한다. 결함을 검출하기 위해서는 이미지 내 특정 영역을 찾아내고 그 영역 안에 있는 특징들을 추출하여 결함 여부를 판단해야 한다. 따라서, 영역을 검출할 수 있는 깊은 학습 기법으로 Mask R-CNN[2]을 적용하고 전처리 과정으로 이미지 필터를 결합하여 비전 검사장비의 결함 검출 능력을 향상시키는 방안을 제안하고 성능을 평가한다.

## 2. 기존 결함 검출 방식

비전 검사장비는 카메라로부터 획득한 영상을 해석하기 용이하도록 이미지 처리하여 결함 검출에 필요한 특징들을 추출하여 결함 여부를 판단한다. 이미지 처리는 입력 영상의 잡음을 제거하고, 밝기 또는 명암 대비를 향상 시킨 후 예지를 강조하게 하고 영상을 보다 더 선명하

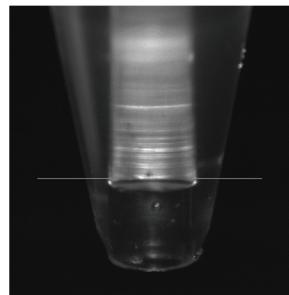


Fig. 1(a) Original image before digital image processing.

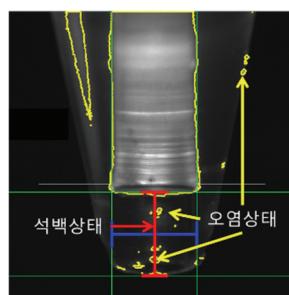


Fig. 1(b) Defect detected by means of the specialized digital image processing.

게 처리한다. 특징 추출은 이미지 처리된 영상의 수직성 분의 픽셀들의 합과 수평성분들의 합을 구하고 적응형 임계 값(Adaptive Threshold)을 적용하는 신호 처리 방식에 의해 구한다. 추출된 특징을 인식하고 해석하여 결함 여부를 판단하는 것은 신경망(Neural Network)으로 구성한다. 포토 공정의 회전 도포 과정에서 발생할 수 있는 결함을 검출하는 방법은 노즐의 석백 상태 및 오염상태, 감광제를 분사하는 시간이다. Fig. 1(b)는 비전 검사장비에서 석백 상태 및 오염 상태를 자동으로 검출한 결과를 보여준다. 비전 검사장비에서 적용하는 이미지 처리를 통한 결함 검출 방식은 회전 도포 과정에서 모터 회전 수 및 분사하는 감광제의 특성에 따라 난반사에 의한 다양한 형태의 영상 왜곡이 발생하고 웨이퍼 종류에 따라 이미지의 밝기(Brightness) 및 대비(Contrast) 또한 여러 형태로 달라진다. 따라서 이미지 처리 및 특징 추출의 신호 처리 방식은 외부 환경요인에 의한 오 검출을 줄이기 위해서는 이미지에 특화된 처리 기법들을 적용해야 한다.

Fig. 2는 비전 검사장비가 이미지 처리 및 신호 처리 방식을 통하여 포토 공정의 회전 도포 과정에서 발생하는 석백 결함을 검출하기 위해 자동으로 석백 특징을 추출한 결과를 보여준다.

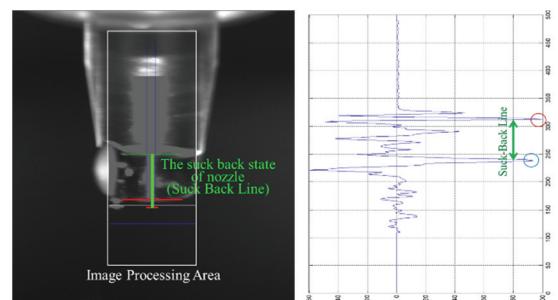


Fig. 2. Stuck-back state extracted by means of using digital image processing and specialized signal processing.

## 3. 새로운 결함 검출 접근법

### 3.1 Mask R-CNN (Regions with Convolution Neural Networks)

외부 환경으로 인해 발생하는 예측할 수 없는 이미지 왜곡 및 다양한 형태의 노즐 특성을 이미지 처리 및 신호 처리 방식으로 해결하기에는 한계가 있다. 이를 극복하기 위해서 이미지 왜곡에도 강인하고 다양한 노즐 형태를 인식하여 그에 맞는 특징을 추출할 수 있는 깊은 학습 방법을 적용한다. 깊은 학습 방법 중 컨볼루션 신경망(CNN, Convolution Neural Networks)은 다양한 이미지를 분류하는데

탁월한 성능을 보이며, 컴퓨터 비전 분야에 많이 활용되고, 반도체 웨이퍼 맵 결합 패턴 분류 및 이미지 검색을 위해서도 컨볼루션 신경망[3]을 사용한다.

주어진 이미지 안에 어느 특정한 클래스에 해당하는 사물이 (만약 있다면) 어느 위치에 포함되어 있는지 ‘픽셀 단위로’ 분할하는 모델을 만드는 것이 세그먼테이션(Segmentation)[4]이다. 시맨틱 세그먼테이션(Semantic Segmentation)은 같은 클래스(Class)인 객체(Object)들은 같은 영역 혹은 같은 색으로 분할하는 것과 달리 인스턴스 세그먼테이션(Instance Segmentation)은 같은 클래스 예측 시에도 객체들을 서로 구별한다.

일반적으로 객체 검출[5-17] 및 이미지 세그먼테이션[18-20]은 ImageNet[21]과 같은 대규모의 이미지 세트로 학습된 모델 가중치를 가지고, 전이학습(Transfer Learning)을 통해 미세 조정하는 연구 결과들이 많이 있다. 하지만, ImageNet[21]에는 반도체 산업분야의 특화된 이미지는 거의 없어 전이학습을 통해서는 결합 검출에 한계가 있으므로 처음부터 학습할 수 있는 네트워크 모델을 구성하여 적용한다.

비전 검사장비에 사용하는 다양한 노즐의 형태는 이미지 분류를 위해 CNN 기반의 ResNet[22]을 적용하고, 결합 검출에 필요한 특징(Feature)들은 Mask R-CNN[2]을 적용하여 찾아낸다. Mask R-CNN[1]은 기존 Faster R-CNN[5]에서 구간 안에 픽셀(Pixel)이 객체인지 아닌지를 판단하는 바이너리(Binary) 마스크(Mask)를 병행하여 추가하고 ROI(Region of Interest) Pooling 대신 ROI Align을 사용한다. ResNet[20] 기반으로 다양한 노즐에 대해 학습한 가중치(Weight)를 Mask R-CNN[2]의 초기 값으로 로드(Load)하여 검사를 위한 특징들을 추출할 수 있도록 학습시킨다.

### 3.2 Mask R-CNN 의 검출 문제

반도체 포토 공정 회전 도포 과정에서 결합 검출을 위해 비전 검사장비로 실시간으로 획득한 이미지에서 추출해야 하는 특징들은 다양한 형태의 노즐 내 석백 영역 및 오염 영역, 감광제를 분사하는 영역들을 찾아내는 것이다. Fig. 3은 결합 검출을 위한 검사 영역을 보여준다. 다양한 형태의 노즐을 거의 99% 이상 완벽하게 분류하는 ResNet[22]의 모델 가중치를 백본 네트워크로 사용하여 Mask R-CNN[2]을 학습시켜도 검사에 필요한 특징들을 추출하기에는 여전히 문제가 있다.

이미지 분류 문제에 있어 구별적인 특징은 매우 중요한 요소이며, 동 클래스 내부의 분산은 작을수록, 타 클래스 간 분산은 클수록 일반적으로 분류문제를 풀기 쉽다 [23]. 노즐 유형에 대한 CNN 특징지도(Feature Map)는 노즐 유형들 간에는 뚜렷하게 구별되지만, 하나의 노즐 유형에



Fig. 3. Inspection region for detecting defect.

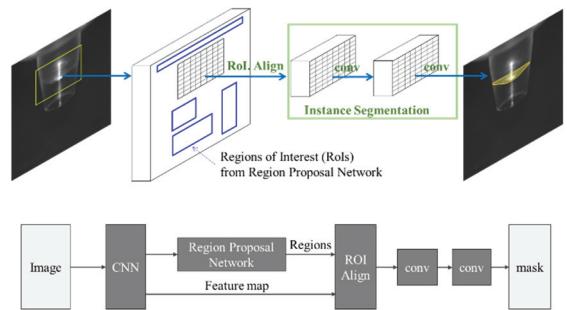


Fig. 4. Instance segmentation process.

대한 CNN 출력을 RPN(Region Proposal Network)의 제안 영역(Proposal Region)들과 매칭하여 특징을 구하면 제안 영역들 간 구별되는 특징을 찾지 못한다. Fig. 4는 인스턴스 세그먼트 처리 과정이다.

제안 영역에서 구별되는 특징을 찾지 못하면 마스크 영역에 오류가 발생한다. 제안 영역에서 구별되는 특징을 찾지 못하는 이유는 그레이 스케일(Gray Scale) 이미지로 해당 영역 안에서 원래의 픽셀이 가지고 있는 정보 만으로는 구별된 특징을 찾기에는 부족하기 때문이다. 이미지가 가지고 있는 원래의 픽셀 정보만으로 구별된 특징지도를 만들기 위해서 다량의 다양한 학습 이미지를 가지고 수 많은 깊은 학습을 수행해도 오 검출을 줄일 수 있을지 모른다. 하지만, 반도체 산업분야에서 반도체 공정에 특화된 이미지를 다량으로 확보하기는 그리 쉽지 않다. 따라서 이를 극복하기 위해 디지털 이미지 처리로 전처리한 구별된 이미지를 깊은 학습의 입력으로 학습하여 원래의 픽셀 정보에서 추출할 수 없었던 구별된 특징지도를 얻을 수 있도록 한다.

### 3.3 전처리와 Mask R-CNN 의 결합

#### 3.3.1 전처리

이미지에서 윤곽선은 이미지의 밝기가 낮은 값에서 높은 값으로 또는 이와 반대로 변하는 지점으로 대체로 객체의 경계(Boundary)로서 모양(Shape), 방향(Direction)을 알 수 있는 여러 가지 정보가 담겨있다. Canny 에지(Edge) 감

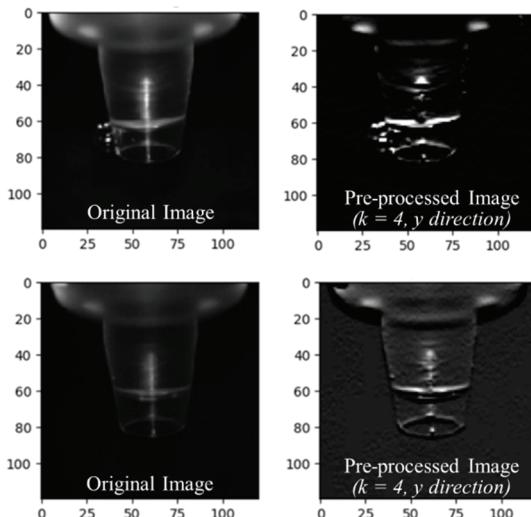
지 알고리즘은 이미지 윤곽을 검출하는 방법 중에 하나이다[24].

검사 유형 중 석백 영역은 노즐 내에 위치하여 노즐 이미지와 노즐 내 감광제 이미지가 겹쳐 있어서 이를 구분하기는 쉽지 않다. 또한 노즐의 특성에 따라 감광제가 비치는 특성이 달라지고 감광제의 종류에 따라 노즐에 비치는 이미지 특성 또한 달라진다. 노즐 특성과 감광제 종류에 따라 이미지 특성은 달라지지만 노즐 내 감광제가 있는 경계선 영역의 특징은 존재한다. 이러한 경계선 영역의 특징은 석백 위치 정보가 될 수 있으므로 이미지에서 윤곽선을 검출하기 위해 소벨(Sobel) 필터[25]를 변형하여 이미지의 밝기에 따라 이득(Gain)을 조정하면서 윤곽선을 검출하는 이미지 필터(Filter)를 적용한다. 원래 이미지에 이미지 필터를 합성곱(Convolution)하여 전 처리를 한다. 이미지에서 밝기 변화율이 크다는 것은 영역의 경계면이 존재한다는 의미로 식 (1), 식 (2)와 같이 수평과 수직 방향으로 편 미분하여 경계면을 구할 수 있다.

$$\frac{\partial f(x, y)}{\partial x} = \Delta x = \frac{f(x + dx, y) - f(x, y)}{dx} \quad (1)$$

$$\frac{\partial f(x, y)}{\partial y} = \Delta y = \frac{f(x, y + dy) - f(x, y)}{dy} \quad (2)$$

수평과 수직 방향으로 두 지점 간의 차이는 픽셀 좌표에서  $dx=dy=1$  임으로 식 (3), 식 (4)와 같이 나타낸다.



**Fig. 5.** The image pre-processed by means of using the specific digital image filter.

$$\Delta x = f(x + 1, y) - f(x, y) \quad (3)$$

$$\Delta y = f(x, y + 1) - f(x, y) \quad (4)$$

윤곽선 검출을 위한 1차 미분 연산자 즉 이미지 필터는 식 (5), 식 (6)과 같다.

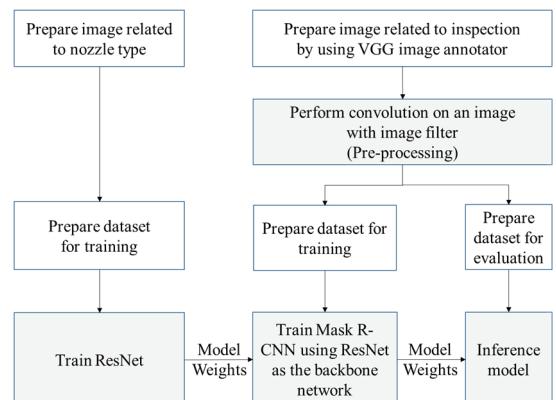
$$\Delta x = \begin{bmatrix} -1 & 0 & 1 \\ k & 0 & k \\ -1 & 0 & 1 \end{bmatrix} \quad (5)$$

$$\Delta y = \begin{bmatrix} -1 & k & 1 \\ 0 & 0 & 0 \\ -1 & k & 1 \end{bmatrix} \quad (6)$$

식 (5), 식 (6)에서  $k$ 는 검출 영역의 이미지 밝기에 따라 적응적으로 가변 적용한다. Fig. 5는 이미지 필터에 의해 전처리된 이미지를 보여준다.

### 3.3.2 전처리를 가진 Mask R-CNN 학습

여러 가지 노즐 유형 이미지를 ResNet[20] 모델로 학습하여 노즐 유형을 분류하게 한다. 노즐 유형에 대해 학습된 ResNet[20] 모델의 가중치를 Mask R-CNN[2] 모델의 초기 가중치 값으로 설정하고 특정 노즐에 대해 석백 영역과 오염 영역, 분사 영역을 검출할 수 있도록 Mask R-CNN[2] 모델을 학습 시킨다. Mask R-CNN[2] 모델을 학습하기 위해 사용하는 이미지는 특정 노즐에 대해 검사 특징들을 가지고 있는 것으로 선별하여 사용하고, 이미지 주석 도구(Image Annotation Tool)인 VGG image annotator 를 사용하여 원본 이미지에서 검출하고자 하는 영역을 설정하고 주석을 만들어서 JSON 파일형태로 저장한다. Mask R-CNN[2] 모델 학습에 사용할 데이터 셋(Dataset)을 만들기 위해 원본 이미지를 불러와서 윤곽선 검출을 위한 이미



**Fig. 6.** Training process of instance segmentation using ResNet as the backbone network.

지 필터와 합성 곱을 적용하여 원본 이미지를 전처리 한다. 전처리된 이미지를 가지고 Mask R-CNN[2] 모델을 학습하기 위한 데이터 셋과 추론을 위한 데이터 셋을 만들어서 Mask R-CNN[2] 모델을 학습시킨다. Fig. 6은 전처리를 가진 Mask R-CNN[2] 모델 학습과정이다.

#### 4. 실험 및 결과

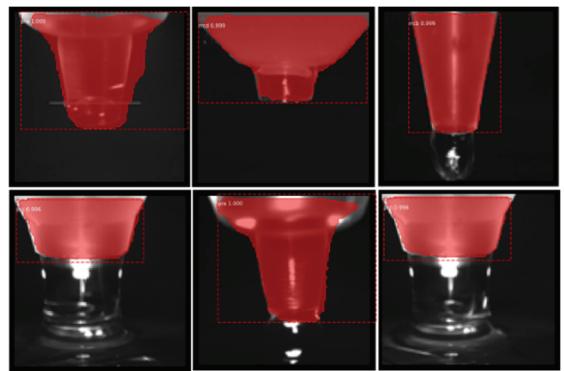
반도체 산업분야의 이미지 특수성 때문에 ImageNet[21]으로 사전 학습된 모델(Pre-trained Model)을 사용할 수 없어서, 처음부터 학습할 수 있는 ResNet[20]을 백본 네트워크 모델로 구성하여 실험하고, Mask R-CNN[1] 모델 학습을 위해 사용하는 하드웨어로는 GPU가 내장된 노트북을 사용한다. Table 1은 노트북에 탑재된 CPU와 GPU에 대한 하드웨어 사양이다.

**Table 1.** Hardware specification

Item	Specification
CPU	Intel Core i7-8750H 2.2GHz
GPU0	Intel UHD Graphics 630
GPU1	NVIDIA GeForce GTX 1050 Ti

비전 검사 장비에서 실시간으로 획득한 이미지는  $680 \times 480$  크기로 그레이 스케일(Gray Scale) 색상을 가진다. 입력 이미지 크기가 커질수록 깊은 학습에 적용될 네트워크 가중치가 많아지므로 짧은 시간 내에 이미지를 처리하고 인식하는데 제한된다. 따라서, 이미지 크기를  $120 \times 120$  크기로 조정하여 ResNet[20] 모델의 입력으로 사용한다. ResNet[20] 모델을 학습하기 위해 사용하는 이미지는 서로 다른 8개의 노즐 이미지, 총 18,304장을 사용한다. ResNet[20] 모델을 학습한 결과로 노즐 이미지 분류 정확도는 99%를 가진다. 노즐 유형을 인식할 수 있는 이미지 학습된 ResNet[20] 모델의 가중치를 Mask R-CNN[2] 모델의 초기 가중치로 적용하고, 노즐 유형별 40장 이미지를 8개 유형에 대해 총 320장의 이미지에 VGG image annotator로 노즐 유형 영역을 설정하여 JSON 파일로 저장한 이미지를 Mask R-CNN[1] 모델의 학습 데이터로 사용한다. Fig. 7은 학습된 Mask R-CNN[1] 모델로 평가한 노즐 유형별 마스크 영역이다.

노즐 유형별 마스크 영역은 전처리 과정이 없어도 마스크 영역을 잘 찾아낸다. 노즐 유형별 마스크 영역 외에 비전 검사에 필요한 노즐의 석백 영역 및 오염 영역, 분사 영역을 검출하기 위해 노즐 유형을 인식할 수 있는,



**Fig. 7.** The experimental result of detecting instance segmentation of nozzle type.

이미 학습된 ResNet[20] 모델의 가중치를 Mask R-CNN[1] 모델의 초기 가중치로 적용한다. 특정 노즐 1개에 대해 266장의 이미지를 VGG image annotator로 석백 영역을 설정하여 JSON파일로 저장한 이미지를 가지고 Mask R-CNN[1] 모델을 학습한다.

전처리 과정이 없이 학습한 경우는 특정 노즐의 석백 영역이 검출 되지 않거나 잘못 인식하는 경우가 있다. Fig. 8은 전처리 과정 없이 학습된 Mask R-CNN[1] 모델로 평가한 특정 노즐의 석백 영역이다.



**Fig. 8.** Suck back region inferred by means of using Mask R-CNN without pre-processing.

비전 검사에 필요한 노즐의 석백 영역 검출을 위해 Mask R-CNN[1] 모델을 학습하는 과정과 학습을 위해 사용한 이미지 파일은 동일하지만, 이미지 필터링을 통한 전처리 과정을 거쳐 학습한 Mask R-CNN[1] 모델과 비교하면, 전처리 과정을 거쳐 학습한 Mask R-CNN[1] 모델이 훨씬 더 좋은 성능을 보인다. Fig. 9는 전처리 과정을 거쳐 학습된 Mask R-CNN[1] 모델로 평가한 특정 노즐의 석백 영역이다.

전처리 과정을 융합한 Mask R-CNN[1] 모델과 전처리 과정이 없는 Mask R-CNN[1] 모델을 각각 동일한 이미지를 가지고 200 epoch까지 학습한다. Table 2는 전처리 과정을 융합한 방식과 단순 Mask R-CNN[1] 모델 방식을 각각 200



**Fig. 9.** Suck back region inferred by means of using Mask R-CNN with pre-processing.

epoch까지 학습하는 과정 중에서 얻은 가장 낮은 손실을 나타낸다. 학습 손실 관점에서는 두 가지 방식이 큰 차이를 보이지 않는다. 따라서, 물체 검출(Object Detection) 알고리즘 성능 평가 방법으로 주로 많이 사용하는 Average Precision(AP)[26]로 두 가지 방식에 대한 성능을 측정하고 비교한다.

**Table 2.** Training loss in the lowest validation loss obtained during the training up to 200 epochs

Dataset	Loss	Pre-processing	
		with	without
Train	loss	0.7468	0.5374
	rpn class loss	0.0148	0.0110
	rpn bbox loss	0.2739	0.1481
	mrcnn class loss	0.0063	0.0036
	mrcnn bbox loss	0.1710	0.1168
	mrcnn mask loss	0.2808	0.2577
Test	loss	0.6669	0.5905
	rpn class loss	0.0128	0.0111
	rpn bbox loss	0.2146	0.1662
	mrcnn class loss	0.0033	0.0030
	mrcnn bbox loss	0.1777	0.1289
	mrcnn mask loss	0.2585	0.2810

두 가지 방식 모두 학습 손실이 가장 낮은 모델 가중치(Weight)를 적용하여 마스크 영역을 추정하고 이를 가지고 평균 AP를 구한다. 학습 손실 관점에서는 두 가지 방식이 큰 차이가 없지만 AP관점에서는 전처리 과정을 융합한 Mask R-CNN[1] 모델 성능이 훨씬 우월하게 나타난다. Table 3은 두 가지 방식에 대한 성능지표를 나타낸다.

**Table 3.** Comparison of performing pre-processing. (IoU: Intersection over Union)

Method	Train (mAP, %)		Test (mAP, %)	
	IoU:0.50	IoU:0.75	IoU:0.50	IoU:0.75
Mask R-CNN Only	48.9	7.5	39.6	7.3
Mask R-CNN with pre-processing	<b>94.4</b>	<b>51.4</b>	<b>90.6</b>	<b>40.5</b>

## 5. 결론 및 향후 연구

반도체 포토 공정에 사용하는 비전 검사장비는 대부분 이미지 처리 방식으로 조그마한 외부 환경 변화에도 검사 오류가 쉽게 발생한다. 반면에 깊은 학습은 이러한 외부 변화에도 이미지 분류에는 강인한 특성이 있다. 하지만 포토 공정에서 검사를 위해 추출하는 특징들은 특정 영역에 국한되고 CNN의 특징 지도로 추출하기에는 한계가 있다. 다량의 여러 가지 형태의 이미지를 가지고 Mask R-CNN[1] 모델을 아무리 많이 학습하여도 마스크 영역 검출에 오류가 발생한다. 따라서, 이미지 처리 방식의 장점과 깊은 학습의 장점을 서로 융합하여 보다 정확한 성능을 가지도록 한다. 비전 검사장비에서 필요한 정확한 마스크 영역 검출을 위해 이미지 처리를 통한 전처리 과정을 Mask R-CNN[1] 모델과 융합하여 학습시키면 마스크 영역의 정확도 성능 측면에서 훨씬 나은 결과를 얻을 수 있다.

향후에는 석백 영역 검출 외 오염 영역, 분사 영역을 동시에 검출할 수 있는 Mask R-CNN[1] 모델을 학습하여 성능을 평가한다. 또한 학습의 방법에 따라 마스크 영역 검출 정확도를 분석하고자 한다. 학습 방법에는 이미지 주석 도구로 세가지 유형의 검출 영역을 설정하여 한 번에 Mask R-CNN[1] 모델을 학습시키는 방법과 개별로 학습시키는 방법, 단계적으로 학습시키는 방법이 있다. 이러한 학습 방식에 따라 검출 정확도 성능에 미치는 영향을 분석하고, 학습된 모델의 가중치 간 연계할 수 있는 방법에 대해 연구하고자 한다. 또한 ResNet[20] 기반의 Mask R-CNN[1] 모델을 DenseNet[27] 기반으로 Mask R-CNN[1] 모델을 수행하여 성능 부분을 비교 분석한다.

## 참고문헌

- Y. K. C. Yeon Ho Chu, “A Deep Learning based IOT Device Recognition System,” *Journal of the Semiconductor & Display Technology*, vol. 18, no. 2, pp. 1-5, 2019.
- K. He, G. Gkioxari, P. Doll’ar, and R. Girshick, “Mask R-CNN,” *Facebook AI Research (FAIR)*, 2018.

3. T. Nakazawa and D. V. Kulkarni, "Wafer Map Defect Pattern Classification and Image Retrieval Using Convolutional Neural Network," *IEEE Transactions on Semiconductor Manufacturing*, vol. 31, no. 2, pp. 309-314, 2018, doi: 10.1109/tsm.2018.2795466.
4. M. Everingham, L. Gool, C. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303-338, 2010, doi: 10.1007/s11263-009-0275-4.
5. R. Girshick, "Fast R-CNN," *Microsoft Research*, 2015.
6. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *arXiv.org*, 2014.
7. J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object Detection via Region-based Fully Convolutional Networks," *arXiv.org*, 2016.
8. W. Liu, D. Anguelov, C. Szegedy, S. Reed, F. Cheng-Yang, and A. Berg, "SSD: Single Shot MultiBox Detector," vol. 9905, ed. Ithaca, 2016.
9. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," vol. 2016-, ed. 2016, pp. 779-788.
10. L. Tsung-Yi, P. Goyal, R. Girshick, H. Kaiming, and P. Dollar, "Focal Loss for Dense Object Detection," vol. 2017-, ed. 2017, pp. 2999-3007.
11. T. Kong, A. Yao, Y. Chen, and F. Sun, "HyperNet: Towards Accurate Region Proposal Generation and Joint Object Detection," *arXiv.org*, 2016.
12. T. Kong, F. Sun, A. Yao, H. Liu, M. Lu, and Y. Chen, "RON: Reverse Connection with Objectness Prior Networks for Object Detection," *arXiv.org*, 2017.
13. S. Bell, C. Zitnick, K. Bala, and R. Girshick, "Inside-Outside Net: Detecting Objects in Context with Skip Pooling and Recurrent Neural Networks," *arXiv.org*, 2015.
14. P. Chao *et al.*, "MegDet: A Large Mini-Batch Object Detector," *arXiv.org*, 2018.
15. B. Singh and L. Davis, "An Analysis of Scale Invariance in Object Detection - SNIP," *arXiv.org*, 2018.
16. Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into High Quality Object Detection," 2017.
17. Y.-H. L. Hyochang Ahn, "A Research of CNN-based Object Detection for Multiple Object Tracking in Image," *Journal of the Semiconductor & Display Technology*, vol. 18, no. 3, pp. 110-114, 2019.
18. E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640-651, 2017, doi: 10.1109/TPAMI.2016.2572683.
19. B. Hariharan, P. Arbeláez, R. Girshick, and J. Malik, "Hypercolumns for Object Segmentation and Fine-grained Localization," *arXiv.org*, 2015.
20. L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille, "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs," *arXiv.org*, 2016.
21. D. Jia, D. Wei, R. Socher, L. Li-Jia, L. Kai, and F.-F. Li, "ImageNet: A large-scale hierarchical image database," ed. 2009, pp. 248-255.
22. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Microsoft Research*, 2015.
23. W.-Y. Kang and B.-T. Zhang, "Image Classification using Convolutional Neural Networks Based on Discriminative Features," *KOREA INFORMATION SCIENCE SOCIETY*, pp. 645-647 (3 pages), 2016.12.
24. A. L. KABADE and D. V. G. Sangam, "Canny edge detection algorithm," *International Journal of Advanced Research in Electronics and Communication Engineering*, vol. 5, no. 5, May 2016.
25. O. R. Vincent and O. Folorunso, "A Descriptive Algorithm for Sobel Image Edge Detection," *Proceedings of Informing Science & IT Education Conference (InSITE) 2009*.
26. J. Revaud, J. Almazan, R. Sampaio de Rezende, and C. de Souza, "Learning with Average Precision: Training Image Retrieval with a Listwise Loss," *arXiv.org*, 2019.
27. G. Huang, Z. Liu, and K. Weinberger, "Densely Connected Convolutional Networks," *arXiv.org*, 2018.

접수일: 2020년 8월 5일, 심사일: 2020년 8월 18일,  
제재확정일: 2020년 8월 27일