

PCB 부품 검출을 위한 Knowledge Distillation 기반 Continual Learning

강수명[†], 정대원^{**}, 이준재^{***}

Knowledge Distillation Based Continual Learning for PCB Part Detection

Su Myung Gang[†], Daewon Chung^{**}, Joon Jae Lee^{***}

ABSTRACT

PCB (Printed Circuit Board) inspection using a deep learning model requires a large amount of data and storage. When the amount of stored data increases, problems such as learning time and insufficient storage space occur. In this study, the existing object detection model is changed to a continual learning model to enable the recognition and classification of PCB components that are constantly increasing. By changing the structure of the object detection model to a knowledge distillation model, we propose a method that allows knowledge distillation of information on existing classified parts while simultaneously learning information on new components. In classification scenario, the transfer learning model result is 75.9%, and the continual learning model proposed in this study shows 90.7%.

Key words: Deep Learning, PCB Inspection, Continual Learning, Knowledge Distillation

1. 서 론

영상을 활용한 PCB(Printed Circuit Board) 검사
의 정확도 개선을 위해 최근 딥러닝(deep learning)
중에서도 영상 인식과 분류가 뛰어난 CNN(Convolutional neural networks)이 모델이 조립 불량이나, 부품의 외관 검사 등에 주로 사용되고 있다[1]. 하지만 딥러닝 모델의 학습 및 적용과 결과는 수집된 데이터의 양과 질에 크게 의존하고 있어 클래스가 증가하는 환경에서는 지속적으로 데이터를 누적해야 한다는 문제가 발생한다. 특히 학습 환경에서 사용한 영상과 실제 적용 환경에서 검사할 데이터의 특징이나 종류

가 다를 경우 모델 적합도가 낮아져 적용성이 떨어지는 문제가 발생한다. PCB 검사에서도 사용된 부품이나 기관의 종류에 따라 학습된 딥러닝 모델의 적합도가 달라지므로 이러한 변화를 고려한 딥러닝 모델이 필요하다[1-3].

조립된 PCB는 필요한 전자기기마다 사용되는 부품의 종류가 다르며 크기와 모양이 달라져 여러 공장에서 동일한 딥러닝 모델을 사용할 수 없다. 또한 전자기기의 소형화에 따라 부품의 크기가 소형화되어 다양성이 높아질 가능성이 있다. 일반적으로 이러한 문제를 해결하기 위해서는 지속적으로 수집한 데이터를 대용량 스토리지(storage)등에 보관하여 주기

※ Corresponding Author: Joon Jae Lee, Address: (42601) 1095, Dalgubeol-daero, Dalseo-gu, Daegu, Republic of Korea, TEL: +82-53-580-6682, FAX: +82-53-580-5165, E-mail: joonlee@kmu.ac.kr

Receipt date: Jun. 16, 2021, Revision date: Jul. 1, 2021
Approval date: Jul. 13, 2021

[†] Dept. of Computer Engineering, Graduate School, Keimyung University (E-mail: smgang.kmu@gmail.com)

^{**} Mathematics Major, Faculty of Basic Sciences, Keimyung University
(E-mail: dwchung@kmu.ac.kr)

^{***} Faculty of Computer Engineering, Keimyung University
※ This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2020R111A3074667)*

적으로 변화된 데이터의 특징까지 딥러닝 모델이 모두 학습한 후 새로 적용해야 한다. 하지만 PCB 검사를 위한 영상 데이터는 매우 고화질로 대용량 저장공간이 필요하며, 영상 데이터의 개수가 많아질수록 딥러닝 모델의 학습시간이 매우 길어지므로 새로 학습된 모델의 배포에 필요한 시간이 늘어나기 때문에 산업적 측면에서 활용성이 떨어진다는 단점이 있다 [1-3].

PCB 검사에서는 검사 정확도가 매우 중요하지만, 생산 속도도 매우 중요하므로 대량의 데이터를 학습하여 모델을 적용할 시에 소요되는 시간을 고려해야 한다. 즉, 새로운 부품을 위해 적용해야 하는 영상 획득 시간과 학습 시간, 및 적용시간 등을 고려할 때 새로 모델을 학습 할 때 기 누적된 데이터의 사용을 최소화하고, 기존에 학습 된 모델을 활용하여 기존 클래스의 분류 정확도를 학습할 수 있게 할 경우 학습 시간의 큰 단축이 될 수 있다 [1-3].

본 연구에서는 분류 객체가 증가할 때 기존 학습된 딥러닝 모델의 가중치 정보를 잇는 치명적 망각(catastrophic forgetting)[4] 문제를 개선하기 위하여 CenterNet[5]을 개선한 지식 증류(knowledge distillation)[4,6] 모델을 제안한다. 제안한 knowledge distillation 모델은 teacher 모델과 student 모델의 결합 구조를 가지며, 이때 teacher 모델의 ResNet 50 백본(backbone)과 up-convolution 부는 동결 상태이다. Student 모델의 백본 역시 동결상태이나 up-convolution부에서는 teacher 모델의 정보를 증류 받아 새로운 클래스의 학습을 가능하게 한다.

지속학습 연구의 경우 시나리오를 가정하여 점차 학습 개체를 늘리는 방법으로 모델을 검증한다. 본 연구에서는 부품의 소형화에 따른 개체 다양성을 고려하므로, 기존 대형부품 8개에서 9개, 12개, 14개로 부품을 점차 늘려가며 정확도를 확인한다. 이때, 기존 8개 부품에 대해서는 10%의 학습 데이터를 사용하며, 신규 학습 대상에 대해서는 기존 8개 부품 학습 때 사용된 양과 같은 개수만을 사용한다. 전이 학습(transfer learning)[7]을 같은 조건에서 함께 수행하여 본 연구에서 제시하는 모델과 결과를 비교한다.

본 연구는 2장에서 관련 연구를 살펴보고, 3장에서 제안하는 모델 및 시나리오에 대한 상세한 설명을 다룬다. 4장에서는 실험과 결과를 확인하고, 마지막으로 5장에서 본 연구의 결론과 향후 연구를 논한다.

2. 관련 연구

2.1 지식 증류(Knowledge distillation) 기법

최근 딥러닝 모델 분야에서 활용하는 대다수의 딥러닝 학습 프레임워크는 사전학습모델(pretrained model)을 제공하고 있다. 이는 대용량 영상을 활용하여 사전적으로 학습된 모델을 사용함으로써 대략적인 특징맵의 가중치를 이미 계산한 것이며 전이학습(transfer learning)이라고 부른다. 일반적으로 전이 학습에서는 저수준 특징맵을 동결(freezing)한 후 고수준 특징맵에서 미세조정을 수행한다[7]. 따라서 전이 학습에 관한 방법을 지속학습(continual learning, life-long learning, incremental learning)을 위하여 사용하는 경우 치명적 망각 문제가 발생할 수 있다. 이전 학습을 통해 만든 딥러닝 모델의 가중치를 새로 학습시킬 모델에 전이한 후, 이전 학습 데이터 소량과 신규 학습 데이터 대량을 넣어 학습시킬 경우 학습데이터 불균형 문제가 발생한다. 즉, 기존 데이터의 특징 추출에 효과적이었던 가중치가 변환될 것이며, 대량의 신규 데이터 특징 추출을 더 잘하는 방향으로 학습되기 때문이다. 이는 기존 학습된 가중치를 망각하게 되는 문제가 된다[4,6,7].

본 연구의 대상인 PCB 부품 분류 등의 환경에서는 지속적으로 발생하는 데이터가 극도로 많으며, 모두 저장하기에는 저장용량 문제가 발생하므로 [1-3]의 연구에서는 주요한 데이터 일부만을 코어셋으로 남기는 연구를 수행한 바 있다.

이러한 경우에 신규 클래스에 대해 코어셋 데이터 개수 이상의 데이터를 넣어 학습을 할 경우 데이터 불균형 문제에 의해 미세조정 과정에서 신규 클래스만을 잘 분류하고, 과거 클래스를 잊어버리는 치명적 망각 문제가 발생한다. 치명적 망각 문제를 해결하기 위한 여러 가지 방법이 있으나, [4]의 연구에서는 지식 증류 방법을 활용하여 이러한 문제를 해결하였다.

지식 증류 방법은 사전 학습된 모델에서 일부의 정보를 새로운 학습을 하는 모델로 전달하는 방법으로서 두개의 모델에 대한 손실을 최소화하는 방식으로 구현된다. 따라서 학습하고자 하는 모델이 가진 손실함수와, teacher-student 모델 간의 손실함수를 더하여 해당 손실함수를 최소화 하는 방향으로 학습된다[4,6].

이때 teacher 모델과 student 모델은 같은 형태이

거나 다를 수 있고, 증류를 위한 레이어는 두 모델 모두 동일하도록 변환하여 사용해야 한다. 학습 과정에서의 저수준 혹은 고수준 특징 추출 레이어 간의 지식 증류를 수행하는 방법과, 최종 클래스 결정 시에 활용되는 softmax 층에서도 지식 증류를 수행하는 방식으로 기존 연구들이 수행되었다. 즉 teacher 모델의 연산하는 과정 및 특징 추출하는 방법을 student 모델이 모방하도록 하는 것이다. 기존 지식 증류 모델의 경우 앙상블 등의 복잡한 모델은 고성능이긴 하나 연산 측면에서 시간적 문제가 있으므로 이를 모방할 수 있는 복잡도가 낮고 연산 속도가 빠른 모델을 만들기 위한 방법으로 제시 되었다[4,6].

이러한 방법을 통해 기 학습에 활용했던 데이터의 사용을 최소화 하고 teacher 모델이 특징을 추출하는 방법만을 전수해주기 때문에 데이터의 저작권이나 초상권 등을 보호하기 위한 방법으로도 활용되고 있다[6].

2.2 CenterNet

객체 검지 모델은 일반적으로 one stage detector, two stage detector의 두 가지 방법으로 분류 가능하며 각 방법의 장단점이 명확하다. One stage detector의 경우 속도가 매우 빠르지만 정확도가 낮으며, two stage detector의 경우 그 반대다. 후자의 검출기는 일반적으로 물체 검지를 위한 앵커(anchor)와 같이 물체의 크기 영역을 사전에 정의하여 전체 영상에 대해 후보지를 사전에 비교하는 방법을 거치기 때문에 속도가 느리다. 또한 one stage detector의 경우 특징맵의 크기를 줄이면서 물체의 검지와 물체의 크기를 대략 지정하기 때문에 정확한 크기 검출은 힘든 문제가 있다. 본 연구에서 활용하고자 하는 CenterNet[3]의 경우 이러한 문제를 해결하는 one stage detector 방법으로서, 객체 중심점과 물체의 크기, 크기를 보정하는 오프셋(offset)을 검출하는 방식으로 물체를 검출한다. 이 모델은 전체적으로 encoder-decoder 형태를 가지며, 일반적인 의미 분할 모델의 기본 구조와 유사하다. 하지만 의미 분할 모델의 가장 끝 레이어가 의미 분할 결과가 나오는 것에 반해, CenterNet의 경우 decoder의 가장 끝 레이어에서 도출되는 특징맵을 활용하여 3가지 레이어를 각기 학습하여 객체 중심점을 찾도록 하는 히트맵(heat map), 물체 크기, 오프셋(offset)을 산정한다[6].

3. 제안한 방법

3.1 제안 모델

본 연구에서는 지속적으로 증가하는 PCB 부품에 대응할 수 있도록 지식 증류 방법으로 CenterNet 모델을 수정한 지속학습 모델을 제안하고자 한다. 본 연구에서 제안한 지식 증류 CenterNet(이하 KD-CenterNet) 모델은 siamese[8 구조를 가지며 Teacher 모델과 student 모델의 encoder 부의 경우 백본으로 모두 ResNet 50[9]을 사용하며, 동결되어 있다. 또한 decoder 부의 경우 teacher 모델은 동결되어 있고, student 모델의 decoder 부만 학습된다. 즉, teacher의 decoder 부분에서 기존 class에 대한 정보를 전달한다.

CenterNet의 기존 손실 함수는 히트맵을 위한 focal loss, 물체 크기와 오프셋을 위한 L_1 loss로 구성되어 있다. 이때 히트맵을 위한 focal loss를 L_k 로 보고, 크기를 위한 L_1 loss를 L_{size} , 오프셋을 위한 L_1 loss를 L_{off} 라고 할 때, 기존 CenterNet의 손실함수 L_{det} 는 수식 (1)과 같다. 이때 λ 값은 크기를 위한 손실에서는 0.1, 오프셋의 경우 1 이다[5].

$$L_{det} = L_k + \lambda_{size} L_{size} + \lambda_{off} L_{off} \quad (1)$$

제안한 모델 구조는 Fig. 1과 같으며, 본 연구에서 제안한 KD-CenterNet의 경우 decoder 부의 첫번째 레이어의 정보만을 증류하여 전달하며, Fig. 1에서 KD_1 를 의미한다. 즉, KD_1 를 최소화 하는 최종 손실 함수는 수식 (2)와 같다.

$$Loss = L_{det} + \lambda L_{KD_1} \quad (2)$$

이때 L_{KD_1} 는 아래 수식 (3)과 같다. 이때 F 는 특징맵, E 는 딥러닝 모델의 인코딩 부, D 는 딥러닝 모델의 디코딩 부, X_n 은 입력 이미지, n 은 배치 이미지의 숫자이다. k 는 디코더 부분에 대한 레이어 순번이며, 여기서는 1번 레이어를 활용하기 때문에 $k=1$ 이다. H' , W' 는 각각 특징맵의 높이, 너비 크기며 C 는 특징 맵의 채널 개수이다.

$$F = \sum D_k(E(X_n)) \quad (3)$$

$$L_{KD_1} = |F_{teacher} - F_{student}|$$

$$L_{KD_k} = \sum_{H'_k W'_k C_k} L_{KD_1}$$

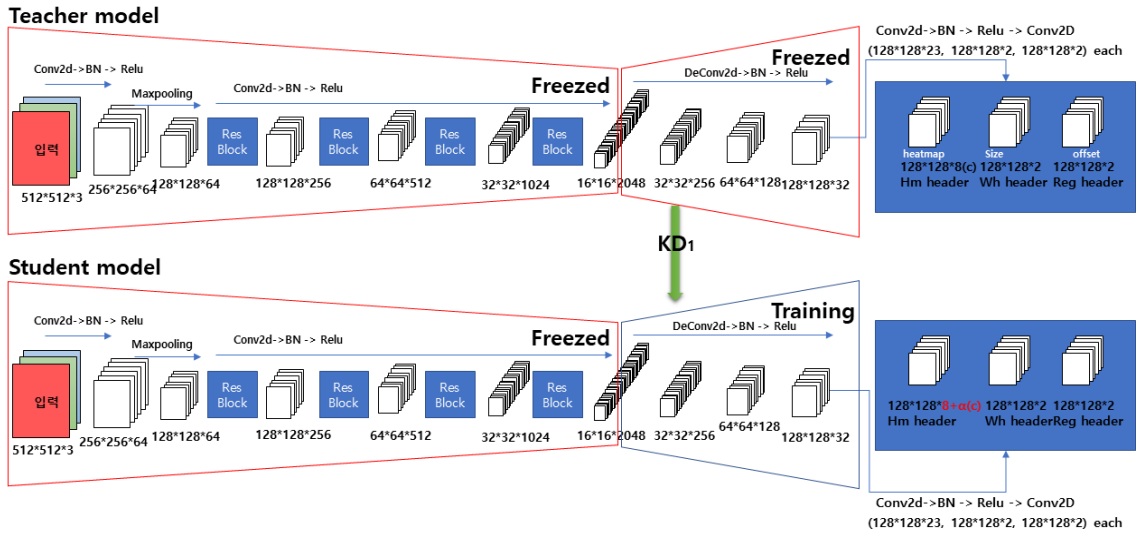


Fig. 1. Proposed KD-CenterNet model architecture.

3.2 시나리오 구성

지속 학습 모델을 검증하기 위해서는 일반적으로 분류하고자 하는 클래스가 증가된 상황을 가정한다. PCB 부품은 트랜지스터, 코일 등의 대분류의 경우 거의 고정되어 있으나 저항 값이나 크기 등의 세부 분류는 증가되는 추세이며, 기기의 소형화 등에 인한 부품 소형화도 부품 증가의 한 이유라고 할 수 있다. 또한 PCB 검사를 수행하는 실제 필드에서도 검사 기간에 따라 서로 다른 PCB종류를 검사해야 하는 경우가 발생하므로, 수집되는 부품의 종류나 불량이 기간에 따라 상이한 경우가 발생한다. 따라서 이러한 상황을 가정하여 지속 학습 모델 학습 시나리오를 구성한다.

우선 1차로 분류하고자 하는 부품은 대분류로 총 3가지이며 세부 부품 분류는 8개이다. 또한 2차로 분류하고자 하는 부품은 대분류 총 4가지이며, 세부 부품 분류의 개수는 총 9개이다. 차례로 3차에서 12개, 4차에서 14개의 부품을 분류한다. 시나리오의 설계 가정은 대분류가 기존 3개에서 4개로 증가하는 상황을 예상하였으며, 3, 4차에서 소분류 중 작은 부품이 점차 증가한다는 것을 반영하였다. 상세 종류는 Table 1과 같다.

이때 세부 분류 이름에 포함된 숫자는 부품의 크기와 같다. 예를 들어 C-CHIP_1608Size의 경우 1608이 크기에 관한 분류명이며 16이 가로, 08이 세로 크기이다. 즉, C-CHIP_0603Size의 가로 06, 세로 03보

다 1608부품이 크다. 앞선 시나리오 가정을 통해 종류가 다양화되고 소형화 됨을 가정하였다. 또한 품의 모양이 전혀 다른 트랜지스터를 2차 분류에서 추가함으로써 부품 모양이 다른 경우의 지속학습에 대한 분류 정확도 변화도 확인하고자 한다.

이때 1차로 일반 CenterNet을 학습시켜 동결 모델을 저장한다. 2차 학습때는 동결시킨 모델을 teacher 모델과 student 모델에 모두 전이시켜 초기화를 한 후, student 모델 decoder부를 제외하고는 모두 동결시킨다. 즉, student의 decoder부만을 학습시킨다.

또한 학습 데이터는 1차와 2차, 3차, 4차에서 차이를 둔다. 2차 학습에서는 1차 학습 때 사용한 데이터를 저장용량 관계로 대폭 삭제함을 가정하여 2차 이상의 실험에서 데이터 불균형을 만든다. 이때 데이터 구성은 각 차수 차례로 Table 2, 3, 4, 5와 같다.

4. 실험 결과 및 고찰

4.1 실험 구성

데이터 구성은 1차와 2차 이상을 달리 구성하였다. 각 클래스마다 학습과 검증의 데이터는 동일하게 구성하였으나 수집된 개수가 달라, 학습과 검증을 동일한 개수로 맞추었기 때문에 테스트를 위한 데이터 개수는 달라질 수 있다.

이때 (2)번 수식에서, 종류 정보가 기존 손실함수보다 큰 영향을 가지도록 손실함수에 적용할 때 λ 값

Table 1. All of experiment learning data classes configuration.

Main category	1 st experiment	2 nd experiment	3 rd experiment	4 th experiment
C-chip (capacitor chip)	C-CHIP_1608Size	C-CHIP_1608Size	C-CHIP_1005Size	C-CHIP_0603Size
			C-CHIP_1608Size	C-CHIP_1005Size
	C-CHIP_2012Size	C-CHIP_2012Size	C-CHIP_2012Size	C-CHIP_2012Size
	C-CHIP_3216Size	C-CHIP_3216Size	C-CHIP_3216Size	C-CHIP_3216Size
Diode	DIODE_1511Size	DIODE_1511Size	DIODE_10065Size	DIODE_10065Size
			DIODE_1511Size	DIODE_1511Size
	DIODE_1810Size	DIODE_1810Size	DIODE_1810Size	DIODE_1810Size
	DIODE_2515Size	DIODE_2515Size	DIODE_2515Size	DIODE_2515Size
R-chip (register chip)	R-CHIP_2012Size	R-CHIP_2012Size	R-CHIP_1005Size	R-CHIP_0603Size
			R-CHIP_2012Size	R-CHIP_1005Size
	R-CHIP_3216Size	R-CHIP_3216Size	R-CHIP_3216Size	R-CHIP_3216Size
Transistor		TR_3PinTR_2812Size	TR_3PinTR_2812Size	TR_3PinTR_2812Size

Table 2. 1st learning data configuration.

Main category	Subcategory	Num. of train	Num. of validation	Num. of Test
C-chip (capacitor chip)	C-CHIP_1608Size	480	120	400
	C-CHIP_2012Size	480	120	1200
	C-CHIP_3216Size	480	120	400
Diode	DIODE_1511Size	480	120	400
	DIODE_1810Size	480	120	107
	DIODE_2515Size	480	120	399
R-chip (registor chip)	R-CHIP_2012Size	480	120	400
	R-CHIP_3216Size	480	120	400

Table 3. 2nd learning data configuration.

Main category	Subcategory	Num. of train	Num. of validation	Num. of Test
C-chip (capacitor chip)	C-CHIP_1608Size	48	12	400
	C-CHIP_2012Size	48	12	1200
	C-CHIP_3216Size	48	12	400
Diode	DIODE_1511Size	48	12	400
	DIODE_1810Size	48	12	1200
	DIODE_2515Size	48	12	400
R-chip (registor chip)	R-CHIP_2012Size	48	12	400
	R-CHIP_3216Size	48	12	400
Transistor	TR_3PinTR_2812Size	480	120	1088

을 4로 두고, teacher 모델은 8개 클래스에서 학습한 모델이다.

학습과 테스트를 위한 학습 환경은 AMD Ryzen 7 2700X Eight-core Processor, 64GB RAM, Ge

Table 4. 3rd learning data configuration.

Main category	Subcategory	Num. of train	Num. of validation	Num. of Test
C-chip (capacitor chip)	C-CHIP_1005Size	480	120	400
	C-CHIP_1608Size	48	12	400
	C-CHIP_2012Size	48	12	1200
	C-CHIP_3216Size	48	12	400
Diode	DIODE_10065Size	480	120	372
	DIODE_1511Size	48	12	400
	DIODE_1810Size	48	12	1200
	DIODE_2515Size	48	12	400
R-chip (registor chip)	R-CHIP_1005Size	480	120	379
	R-CHIP_2012Size	48	12	400
	R-CHIP_3216Size	48	12	400
Transistor	TR_3PinTR_2812Size	480	120	1088

Table 5. 4th learning data configuration.

Main category	Subcategory	Num. of train	Num. of validation	Num. of Test
C-chip (capacitor chip)	C-CHIP_0603Size	480	120	385
	C-CHIP_1005Size	480	120	400
	C-CHIP_1608Size	48	12	400
	C-CHIP_2012Size	48	12	1200
	C-CHIP_3216Size	48	12	400
Diode	DIODE_10065Size	480	120	372
	DIODE_1511Size	48	12	400
	DIODE_1810Size	48	12	1200
	DIODE_2515Size	48	12	400
R-chip (registor chip)	R-CHIP_0603Size	480	120	400
	R-CHIP_1005Size	480	120	379
	R-CHIP_2012Size	48	12	400
	R-CHIP_3216Size	48	12	400
Transistor	TR_3PinTR_2812Size	480	120	1088

Force RTX 2080 그래픽카드를 사용하였으며 딥러닝 프레임워크는 python 3.6, Anaconda 3, Tensorflow 1.13.1 backend Keras 2.2.4-tf와 같이 구성하였다. 사용한 영상은 512×512×3의 영상을 사용하였다.

4.2 실험 결과

우선 시나리오 기반의 실험을 검증하기 이전에 일반적인 상황에서의 8개 클래스와 14개 클래스에 대한 학습 실험을 수행하였다. 각 클래스별 상세 결과는 Table 6에서 기술 하였으며 각 결과는 mAP(mean

Average Precision)로 표기하였다.

모든 클래스에 대해 학습 480장, 검증 120장을 사용하여 데이터 균형을 맞추었다. 이때 결과를 보면 보편적으로 작은 부품을 인식하는 것이 어려운 것으로 보인다. Fig. 2는 부품 데이터의 예시이며 타 부품에 비해 0603 사이즈 부품이 매우 작은 것을 확인할 수 있다.

앞선 Table 6에서의 학습 모델 중 1차 8클래스에 대한 모델을 teacher 모델로 두고 전이학습과 지속학습을 수행하였다. Table 7과 8에서는 전이학습과 제

Table 6. Experimental results in general learning situations (mAP).

Experiment		8 classes	9 classes	12 classes	14 classes
C-CHIP_0603Size		-	-	-	0.93
C-CHIP_1005Size		-	-	0.99	0.98
C-CHIP_1608Size		1	1	0.99	1
C-CHIP_2012Size		1	1	1	1
C-CHIP_3216Size		1	1	1	1
DIODE_10065Size		-	-	1	1
DIODE_1511Size		0.99	0.99	1	1
DIODE_1810Size		1	1	1	1
DIODE_2515Size		1	1	1	1
R-CHIP_0603Size		-	-	-	1
R-CHIP_1005Size		-	-	0.98	0.98
R-CHIP_2012Size		1	0.99	1	1
R-CHIP_3216Size		1	0.99	1	0.98
TR_3PinTR_2812Size		-	1	1	1
Average accuracy	IOU 0.2	0.998	0.998	0.996	0.994
	IOU 0.5	0.998	0.998	0.996	0.994

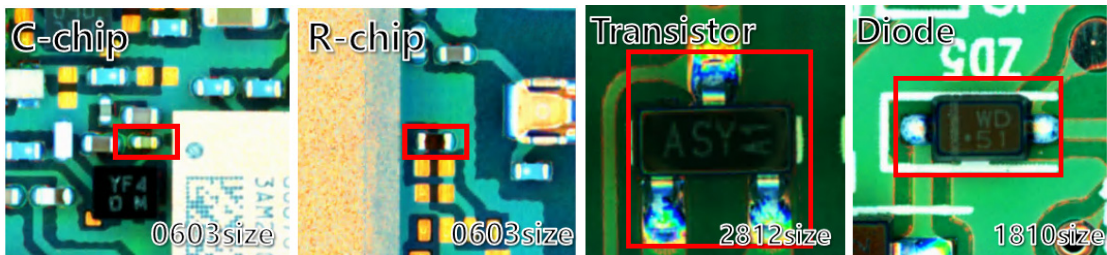


Fig. 2. Example of components data.

안한 지속학습 모델을 검증하였다. 전이학습 결과에서는 기존 클래스는 어느 정도 잘 분류하는 것으로 보이나 신규로 들어온 클래스에 대해서는 전혀 인식 못하는 경우도 종종 보인다. 신규로 들어온 클래스의 문제로 인해 기존 클래스에 대해서도 일부 망각의 문제가 발생된 것이 보인다. 특히 C-CHIP_1608 등의 경우가 치명적 망각 문제를 일으킨 경우이다. 학습 시 충분한 데이터가 없는 경우 기존 클래스 분류 정확도가 낮아진다.

특히 전이학습 4회차의 작은 부품(C-CHIP_0603 Size, R-CHIP_0603Size)에 대해서는 전혀 학습이 제대로 이루어지지 않고 있으며, 기존 8 클래스(1차)에서는 잘 인식하던 C-CHIP_1608Size 부품 등도 3차, 4차 실험에서 mAP(mean Average Precision)가

0.08, 0.15로 떨어지는 것을 볼 때 전이학습 중 기존 모델의 정보를 잊어버리는 치명적 망각 문제가 발생한다고 할 수 있다.

R-CHIP_2012Size 부품의 경우 전체적으로 mAP가 1인 것을 고려할 때 모든 클래스에 대한 망각 문제가 있는 것은 아니며, 특징 추출이 상대적으로 쉽고, 특이점이 있는 경우 정확도가 보존된다는 것도 확인할 수 있다. Fig. 3의 (a)와 (b)를 비교해볼 때 (b)의 R-CHIP_2012Size 부품은 상단에 숫자가 작성되어 있어 다른 부품에 비해 특징이 뚜렷하기 때문에 망각이 없이 mAP가 보존된 것으로 판단된다. 하지만 크기 외에 큰 특징이 없는 경우 신규 데이터로 인지하는 경우가 발생하며, 이는 전이학습 4차 실험의 복합 행렬 Fig. 4에서 확인할 수 있다.

Table 7. Results from transfer learning experiments (mAP).

Experiment		2 nd experiment	3 rd experiment	4 th experiment
C-CHIP_0603Size		-	-	0
C-CHIP_1005Size		-	0.55	0.55
C-CHIP_1608Size		1	0.08	0.15
C-CHIP_2012Size		1	0.96	0.99
C-CHIP_3216Size		1	1	1
DIODE_10065Size		-	0.64	0.7
DIODE_1511Size		0.84	0.7	0.72
DIODE_1810Size		1	1	1
DIODE_2515Size		0.96	0.98	0.98
R-CHIP_0603Size		-	-	0
R-CHIP_1005Size		-	0.4	0.22
R-CHIP_2012Size		1	1	1
R-CHIP_3216Size		0.96	0.94	0.95
TR_3PinTR_2812Size		0.98	0.96	0.96
Average accuracy	IOU 0.2	0.981	0.833	0.759
	IOU 0.5	0.981	0.831	0.758

Table 8. Results of continual learning experiments using the proposed model KD-CenterNet (mAP).

Experiment		2 nd experiment	3 rd experiment	4 th experiment
C-CHIP_0603Size		-	-	0.63
C-CHIP_1005Size		-	0.95	0.92
C-CHIP_1608Size		0.99	0.49	0.74
C-CHIP_2012Size		0.98	0.99	0.98
C-CHIP_3216Size		1	1	0.99
DIODE_10065Size		-	0.98	0.97
DIODE_1511Size		0.97	0.7	0.7
DIODE_1810Size		1	1	1
DIODE_2515Size		0.98	0.99	0.99
R-CHIP_0603Size		-	-	0.57
R-CHIP_1005Size		-	0.83	0.8
R-CHIP_2012Size		1	0.93	0.97
R-CHIP_3216Size		0.96	0.94	0.98
TR_3PinTR_2812Size		1	1	1
Average accuracy	IOU 0.2	0.985	0.93	0.907
	IOU 0.5	0.985	0.929	0.902

Fig. 4의 R-CHIP_1608Size 부품의 대다수 예측 결과는 R-CHIP_1005Size로 인식하는 것을 알 수 있다. TR_3PinTR_2812Size의 부품은 Fig. 2를 볼 때 가장 뚜렷하게 다른 모양을 보여주므로 단독적으로

학습이 잘 된 것을 알 수 있으며, 다른 부품의 망각에도 영향을 주지 않는다. 하지만 레이블 순서로 볼 때 가장 가까이 있는 R-CHIP_3216Size부품 인식에 영향을 주는 것을 확인 할 수 있다. 이는 최종적으로

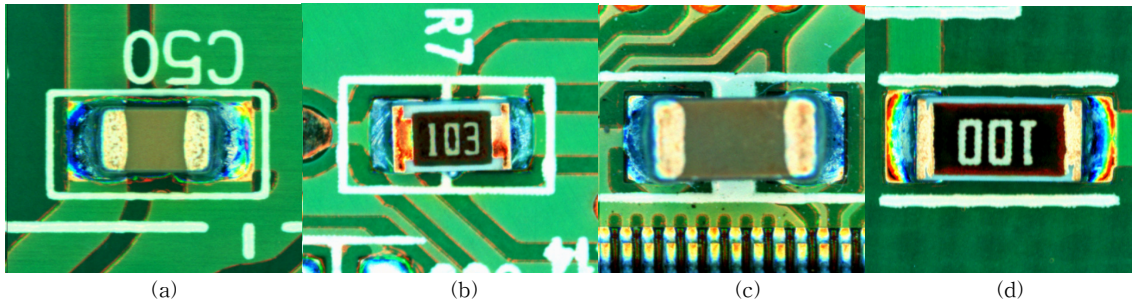


Fig. 3. Examples of 2012 and 3216 sized components. (a) C-CHIP_2012Size, (b) R-CHIP_2012Size, (c) C-CHIP_3216Size, and (d) R-CHIP_3216Size.

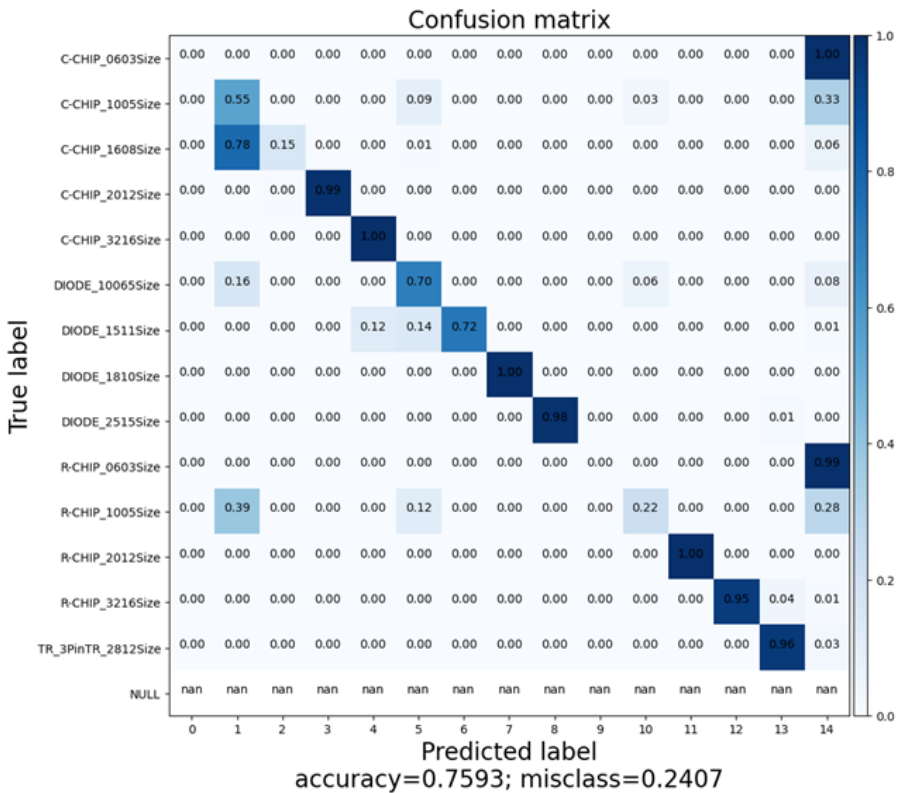


Fig. 4. Confusion matrix of 4th transfer learning result.

부품의 클래스를 결정 할 때 가장 가까운 레이어의 가중치에 영향을 준다는 것을 알 수 있다.

본 연구에서 제안한 모델의 경우 기존 클래스에 대해 일부 망각이 발생하나 전이학습보다는 나은 결과를 보인다. 또한 4차 학습에서의 mAP 평균 결과를 볼 때, 지속 학습 모델에서 정확도로 환산할 시 14.4% 이상 좋은 결과를 보이므로 전이학습 대비 지속 학습 모델이 신뢰성이 높다고 할 수 있다. Fig. 5는

제안된 모델 4차 학습 실험의 복잡행렬 결과이다.

Fig. 5의 복잡행렬 결과를 보면 이전 Fig. 4와 달리 기존 8개의 클래스에 대한 치명적 망각이 줄어든 것을 알 수 있다. 학습 데이터의 불균형이 있더라도 제안된 모델의 지식 증류 방법을 활용하면 기존 학습 모델의 판단 방법을 모방하고 가중치를 상당 부분 보존할 수 있다. 그 중에서도 가장 눈에 띄는 부품은 기존에 존재했던 DIODE_1511Size의 부품이며 새로

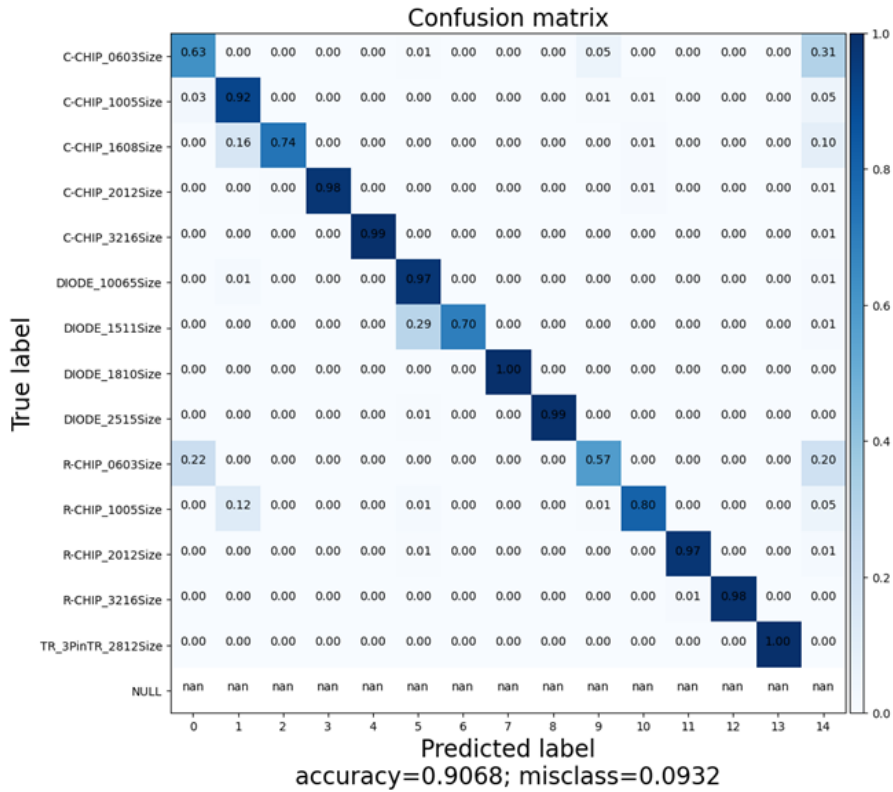


Fig. 5. Confusion matrix of 4th proposed continual learning result.

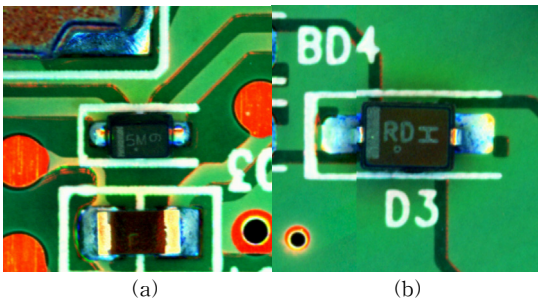


Fig. 6. Example of DIODE Components. (a) DIODE_10065Size and (b) DIODE_1511Size.

추가된 DIODE_10065Size로 오인식 하는 것을 알 수 있다. 이는 데이터 불균형 문제로 학습 과정에서 크기 면에서는 상당히 다르지만 모양이 유사한 부품 (Fig. 6)에 영향을 받는 것으로 추측된다. 이는 C-CHIP_1608Size부품에서도 일부 C-CHIP_1005Size로 오인식 하는 것을 볼 때, 치명적 망각 원인은 학습 과정에서 인근 레이어에 영향을 받는 것과 학습 시 특징 추출에 혼란을 주는 검지 대상의 유사한 모양이

라고 할 수 있다. 또한 작은 부품의 신규 학습에 대해서는 기존 학습 방법 대비 정확도가 매우 낮으므로 이를 해결할 방법이 필요해 보인다.

5. 결론

본 논문에서는 물체 검지가 가능한 CenterNet을 지식 증류기반 모델로 변경하여 지식 증류 손실함수를 통해 새로 증가한 클래스에 대해 학습하면서 기존 클래스에 대해 치명적인 망각을 최소화하는 지속 학습 모델을 제안하였다. 일반적인 상황에서 14개 클래스의 학습 정확도는 IOU 0.2에서 mAP 기준 0.994, 전이학습에서 0.759, 본 논문에서 제안한 지속학습 모델에서는 0.907로서 모든 데이터를 다 저장하는 방법에 비해서는 낮은 수치이나 단순 전이학습에 비해서는 높은 정확도를 보인다는 것을 알 수 있다.

기존 클래스의 망각은 최소화 하지만, 새로운 특징 추출이 필요한 작은 부품이 추가될 때는 이 방법이 효과적이지 않다. 또한 일반적인 학습에서의 정

확도에 비해 지속학습 모델의 한계가 있으므로 향후 이를 개선하는 모델과 손실함수 제안이 필요하다. 또한 작은 부품에 대해 해상도 조절이나, 작은 부품의 정보를 유지할 수 있는 알고리즘의 적용을 수행할 예정이다.

그뿐만 아니라, 기존 데이터의 일부를 저장하는 대신 기존 데이터 전체를 삭제하고 신규데이터와 학습된 모델만으로 학습할 수 있는 지속학습 모델 개발이 필요하다. 이러한 방법을 통해 신속한 딥러닝 모델 학습 및 적용이 가능하여 PCB 생산에 효율성을 높일 수 있다.

REFERENCE

- [1] S. Gang and J. Lee, "Coreset Construction for Character Recognition of PCB Components Based on Deep Learning," *Journal of Korea Multimedia Society*, Vol. 24, No. 3, pp. 382-395, 2021.
- [2] S. Gang, N. Fabrice, and J. Lee, "Coresets for PCB Character Recognition based on Deep Learning," *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, Fukuoka, Japan, 2020, pp. 637-642, 2020.
- [3] S. Gang, N. Fabrice, D. Chung, and J. Lee, "Character Recognition of Components Mount ed on Printed Circuit Board Using Deep Learning," *Sensors*, Vol. 21 No. 9, 2921, 2021.
- [4] U. Michieli and Z. Pietro, "Incremental Learning Techniques for Semantic Segmentation," *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019.
- [5] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as Points," *arXiv preprint*, arXiv:1904.07850, 2019.
- [6] G. Hinton, V. Oriol, and D. Jeff. "Distilling the Knowledge in a Neural Network," *arXiv preprint*, arXiv:1503.02531, 2015.
- [7] S. Park and D. Kim, "Comparison of Image Classification Performance in Convolutional Neural Network according to Transfer Learning," *Journal of Korea Multimedia Society*, Vol. 21 No. 12, pp. 1387-1395.
- [8] P. Gao, R. Yuan, F. Wang, L. Xiao, H. Fujita, and Y. Zhang, "Siamese Attentional Keypoint Network for High Performance Visual Tracking," *Knowledge-Based Systems*, Vol. 193, pp. 1-29, 2020.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *arXiv Preprint*, arXiv:1512.03385, 2015.



강 수 명

2011년 계명대학교 게임모바일콘
텐츠학과 졸업 (학사)
2013년 계명대학교 미디어아트학
과 게임모바일전공 졸업
(석사)
2014년~2017년 ㈜지오씨엔아이
공간정보기술연구소 선임
연구원

2016년~현재 계명대학교 컴퓨터공학과 박사과정
관심분야 : 영상처리, 패턴인식, 컴퓨터비전, 딥 러닝



정 대 원

2002년 중앙대학교 수학과 졸업
(학사)
2004년 중앙대학교 수학과 졸업
(석사)
2010년 University of New
Mexico 졸업(박사)

2010년~2012년 University of New Mexico 겸임강사
2012년~2014년 인하대학교 응용수학연구소 박사 후 연
구원
2014년~현재 계명대학교 기초과학부 조교수
관심분야 : 조합해석학, 편미분방정식, 역문제, 인공지능



이 준 재

1986년 경북대학교 전자공학과
졸업(학사)
1990년 경북대학교 전자공학과
졸업(석사)
1994년 경북대학교 전자공학과
졸업(박사)

1995년~2007년 동서대학교 컴퓨터 정보공학부 부교수
1998년~1999년 Georgia Institute of Technology 방문
교수
2000년~2001년 (주)파미 연구소장
2007년~현재 계명대학교 컴퓨터공학부 교수
관심분야 : 영상처리, 3D 컴퓨터 비전, 딥 러닝, 게임