

방사선 투과 이미지에서의 용접 결함 검출을 위한 딥러닝 알고리즘 비교 연구

Comparative Study of Deep Learning Algorithm for Detection of Welding Defects in Radiographic Images

오상진¹, 윤광호¹, 임채욱¹, 신성철^{1*}

Sang-jin Oh¹, Gwang-ho Yun¹, Chaeg Lim¹, Sung-chul Shin^{1*}

〈Abstract〉

An automated system is needed for the effectiveness of non-destructive testing. In order to utilize the radiographic testing data accumulated in the film, the types of welding defects were classified into 9 and the shape of defects were analyzed. Data was preprocessed to use deep learning with high performance in image classification, and a combination of one-stage/two-stage method and convolutional neural networks/Transformer backbone was compared to confirm a model suitable for welding defect detection. The combination of two-stage, which can learn step-by-step, and deep-layered CNN backbone, showed the best performance with mean average precision 0.868.

Keywords : Radiographic Testing, Welding Defect, Deep Learning

¹ 부산대학교 조선해양공학과

^{1*} 교신저자, 부산대학교 조선해양공학과, 교수
E-mail: scshin@pusan.ac.kr

¹ Corresponding Author, Dept. of Naval Architecture and Ocean Engineering, Pusan National University

^{1*} Dept. of Naval Architecture and Ocean Engineering, Pusan National University

1. 서론

용접공정은 조선·해양 분야에서 전체 공정 중 60% 이상을 차지하고 있다[1]. 용접부에 대한 검사는 비파괴검사로 방사선 투과검사, 초음파 탐상검사, 자분 탐상검사 등 다양한 기술이 사용되고 있으며 특히, 그중에서도 선박의 선주사는 다른 비파괴검사보다 영구적으로 보관할 수 있고, 모든 재료의 용접부 내부를 육안으로 확인할 수 있는 방사선 투과검사를 선호하고 있다.

현재 국내외 조선소에서는 용접공정을 검사하기 위해 많게는 500블록 이상의 구조물을 사람이 직접 용접 검사를 수행하고 있다. 1블록당 2,000여 곳 이상의 용접정보를 수동으로 작성하기 때문에 누락 및 오작이 발생하여 추가 작업이 이루어져 많은 시간과 비용이 요구된다. 수동적으로 행해지는 검사를 일관적이고, 합리적인 결과로 도출하기 위해 검사관의 이해를 돕는 검사의 자동화와 객관화 시스템이 필요한 실정이다.

용접부 결함을 자동으로 검출하기 위한 연구들은 이전부터 이루어지고 있었고, 그중에서 [2]는 특징을 추출하기 위한 방법으로 이미지 전처리를 사용하고, 자동으로 판별하는 부분은 신경망을 이용하여 용접 결함의 종류를 분류했다. 여기서 사용된 신경망의 형태는 multi-layer perceptron (MLP)으로, 결함을 분류하는 용도로만 사용되어 이미지를 분류하고 판독하는 신경망이라고 보기 어렵다. 최근 딥러닝 중에서 이미지 분류로 많은 연구가 이루어지고 있는 convolutional neural networks (CNN)는 기존 알고리즘과 비교하여 높은 성능을 보인다. [3]에서는 CNN을 이용하여 기존에 사용되었던 HOG[4]나 SIFT[5]보다 높은 성능을 보였으며, [6]에서는 CNN과 SVM을 이용하여 신경망에서 추출한 이미지 특징을 기반으로 결함의 영역을 분류하는 연구를 수행했다.

하지만 용접 결함은 결함의 종류에 따라 적용되는 규정이 달라지기 때문에 결함의 영역뿐만 아니라 결함의 종류 분류도 검사 자동화에서 필수적으로 수반되어야 한다. 하나의 신경망에서 객체의 종류 분류와 위치 회귀가 가능하도록 객체 검출 방법이 Fast R-CNN[6], Faster R-CNN[7]으로 발전되어 객체 검출에 주로 사용되고 있다. 이미 방사선 투과 이미지와 같은 방사선 이미지를 다루는 의료 분야에서는 CNN의 기법으로 객체를 검출하는 연구가 이루어지고 있으며, 초음파 이미지에서 CNN을 이용하여 신경 영역을 검출하는 연구도 이뤄졌다[7]. 방사선 이미지에서 객체를 검출하는 성능은 여러 연구에서 확인할 수 있으며, 이를 통해 방사선 투과 용접 이미지에서도 CNN 기법의 적용 가능성을 확인할 수 있었다.

최근 자연어 처리 영역에서 성과를 보인 Transformer[8] 구조가 이미지 인식 분야에 적용되면서 각종 인식 대회 및 연구 성과에서 높은 지표를 보인다. Transformer 구조는 입력 시퀀스(sequence)를 하나의 벡터로 압축하는 과정에서 일부 정보가 손실되지 않는다는 장점과 이미지를 패치로 나눠 토큰 형식으로 이미지 전체를 학습할 수 있다는 장점이 어우러져 낮은 계산 비용과 높은 성능으로 연결된다[9]. 이미지 인식에 대한 전반적인 가능성과 용접 결함의 특징에서 볼 수 있는 작은 객체에 대한 탐지에 대한 성능[10], 소규모 데이터 세트에서의 성능[11]을 확인하고 Transformer를 이용하여 용접 결함을 자동으로 검출하기 위한 연구를 수행했다.

본 논문에서는 객체 검출 정확도에서 높은 성과를 보인 Transformer와 CNN을 이용하여 방사선 투과 이미지에서 용접부 결함을 자동으로 검출하는 알고리즘을 제안한다. Swin transformer[12]와 ResNet[13]을 backbone 네트워크로 구성하여 비교했으며, Retinanet[14], Faster R-CNN, Cascade

R-CNN[15]을 객체 탐지 모델로 설정하여 one-stage와 two-stage에 대한 성능도 분석했다.

2. 관련 연구

ASME Section V에 따라 용접 결함은 11종류로 나뉘며, 비파괴 검사 중 방사선 투과검사(Radiographic Testing, RT)로 확인할 수 있는 결함은 10종류로 분류된다. 검사의 특성상 육안으로 검사가 가능한 결함의 경우, 방사선 투과검사 이전에 보강이나 재작업을 수행하기 때문에 필름 수가

적을 뿐만 아니라, 산업체의 보안과 맞물려 데이터 확보가 어려운 실정이다. 따라서 Table 1과 같이 방사선 투과검사로 확인할 수 있는 결함을 10종류에서 11종류로 세분화하고, 육안 검사로 제외되는 결함 burn through, inadequate reinforcement를 제외하여 9종류의 결함을 분류 대상으로 선정했다.

방사선 투과검사는 탐상면에 평행하게 위치하는 slag inclusion, porosity 등의 체적 결함에 대해서는 검출 성능이 우수하지만, 수직으로 나타나는 결함인 incomplete fusion, crack 등의 선형결함에 대해서는 검출하기 어려운 단점이 있다[16]. 기하학적 형상 및 방향성을 가지는 체적 결함과 선형결함을 검출하기 위해서는 데이터 기반 패턴 분석 및 검출 가능한 알고리즘이 요구된다. 방사선 투과검사 아날로그 필름은 스캐너를 이용하여 전환하면 고분해능 디지털 해상도인 4,096픽셀까지 표현할 수 있다[17]. 이때 용접 결함의 크기를 측정하면 Fig. 1과 같이, 100² 픽셀 미만 크기가 53% 이상 구성되어 있으며, Common Object in Context (COCO)[18]에 따라 32² 미만은 small, 32² ≤ Area < 96²은 medium 객체로 분류되어 비교적 작은 결함에 속하게 된다. 이미지 크기에 비해 작은 결함의 패턴을 분석하고, 검출하기 위해 이미지 인식 분야에서는 딥러닝 모델로서 one-stage와 two-stage 방법이 특징에 따라 연구되고 있다.

딥러닝에서의 객체 검출(object detection)은 후보 영역 탐색(regional proposal)과 분류(classification)로 구성되며, 수행되는 과정을 구분함에 따라 one-stage와 two-stage 방법으로 나뉜다. two-stage로 구성된 Faster R-CNN은 후보영역을 탐색하는 과정을 학습할 수 있는 모델인 region proposal network (RPN)로 구성하여 기존 알고리즘보다 객체 검출 정확도를 향상하였다.

Table 1. Imperfections detection according to the type of NDE method. Dot: All or most standard techniques will detect this imperfection under all or most conditions. Empty dot: One or more standard technique will detect this imperfection under certain conditions.

Welding Imperfections	Visual Test	Radiography Test	Dataset Classes
Burn Through	●	●	
Cracks		○	
Excessive Reinforcement	●	●	
Inadequate Reinforcement	●	●	
Inclusion (Slag)		●	
Inclusion (Tungsten)		●	
Incomplete Fusion	○	○	
Incomplete Penetration	○	●	
Porosity	●	●	
Root Concavity	●	●	
Undercut	●	●	

RPN에서 후보 영역의 위치를 결정하는 anchor의 크기와 개수를 검출하고자 하는 객체의 크기에 맞게 변형하면 사용자화된 객체의 검출 정확도를 향상할 수 있다[19]. Cascade R-CNN에서는 더 정밀한 객체 탐지를 위해 분류 모듈을 여러 단계로 쌓아 intersection over union (IoU) 기준을 0.5, 0.6, 0.7로 높이면서 학습을 진행한다. 그 결과 multi-stage 객체 검출 방법이 분류 성능뿐만 아니라 위치 검출 성능까지 높일 수 있었다.

One-stage는 한 모듈에서 후보 영역 탐색과 분

류를 동시에 처리하여 계산 효율이 높고, two-stage 방법보다 빠른 장점이 있지만, 정확도가 낮은 한계가 있다. 학습 시 객체와 배경 클래스 불균형으로 인해 가중치가 편향되는 현상이 발생하는데, one-stage는 전체 이미지를 모두 탐색하는 dense sampling 방법을 사용하기 때문에 클래스 불균형이 크게 발생하게 된다. Retinanet에서는 불균형을 해결하고자 Focal loss를 사용하여 정확도를 향상하였다. 객체 존재 여부에 따라 구분되는 positive/negative 후보 영역에 가중치를 두어 균형을 주고, 분류 난이도를 easy/hard 나누어 hard negative 후보 영역에 대해 집중학습하는 손실 함수로서 해당 모델뿐만 아니라 다른 모델에서도 좋은 성능을 보인다[20].

객체 검출 모델은 검출 방법뿐만 아니라 이미지의 특징을 추출하고 분류하는 뼈대(backbone)의 구성 방법에 따라 성능의 차이가 발생한다[21]. 주로 backbone 네트워크는 ResNet을 사용하여 층을 깊게 구성하고, 사전 학습된 가중치를 불러들여 전이학습을 수행하게 되는데, ImageNet[22], COCO dataset 등 많은 데이터가 학습된 가중치를 활용한다는 점에서 정확도와 편의성이 높은 장점이 있다. ResNet의 특징으로 기존 신경망에 shortcut을 연결하여 잔차(residual)를 최소화하는 방향으로 학습된다. 신경망이 깊어질수록 미분하는 횟수가 많아져 출력에 영향을 끼치는 가중치가 작아지는 현상인 vanishing gradient 문제를 잔차를 적용함으로써 해결했다. ResNet과 같이 대부분의 backbone 네트워크는 CNN을 기반으로 구성되어 있다.

자연어 처리에서 활발히 사용되었던 Transformer 구조가 이미지 분야에서도 적용되면서 분류, 객체 검출, 객체 분할 성능에서 기존 CNN 기반 모델을 앞서고 있다. ViT[9]는 이미지를 패치로 분할 후 시퀀스로 변환된 값이 self-attention의 입력으로 사용된다. 입력은 쿼리(query)와 키(key) 벡터

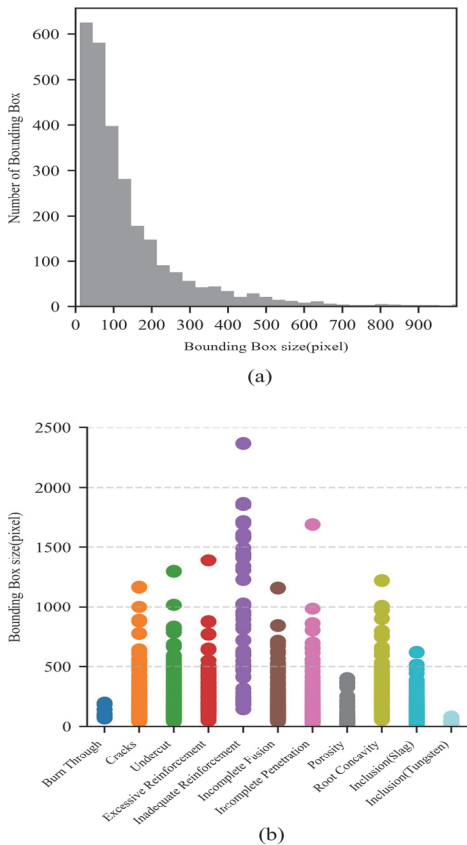


Fig. 1 Defect (bounding box) size in the dataset. (a) Number of bounding boxes by size (square root of area). (b) Distribution of bounding box size by defect

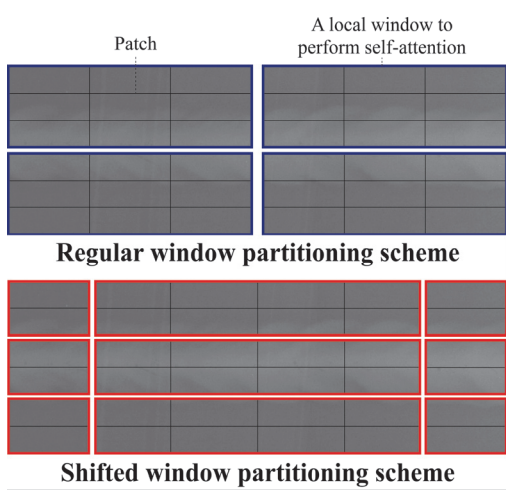


Fig. 2 An illustration of the shifted window approach for computing self-attention in the proposed Swin transformer architecture

로 임베딩되고, 가중치를 계산하여 값(value) 벡터로 산출된다. self-attention을 다중으로 설정하고, MLP 블록을 차례로 연결하여 최종적으로 이미지의 클래스를 분류한다. 계산의 효율성과 모델의 확장성이 기존 CNN 기반 모델보다 좋지만, 우수한 성능을 발휘하기 위해서는 대규모 학습 데이터가 필요하며, 고정된 크기의 패치 사용으로 인해 고해상도 이미지에는 적합하지 않다는 단점이 있다.

Fig. 2와 같이 Swin transformer에서는 shifted windows 분할 방식을 사용하여 window 내에 있는 패치만 self-attention을 수행하여 기존보다 연산량을 줄일 수 있었다. 이에 따라 파라미터 수가 줄어들고, 기존 대비 필요 학습 데이터 수가 적어도 효율적인 학습이 가능해진다.

3. 연구 방법

본 연구에서는 방사선 투과 이미지에서의 용접

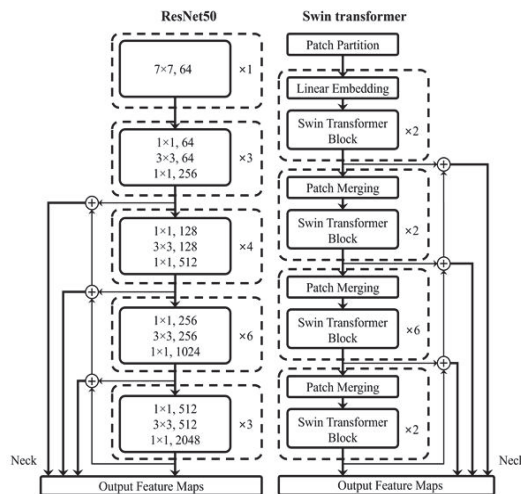


Fig. 3 The architecture of the backbone networks ResNet (left) and Swin transformer (right). ResNet 50 consists of 5 blocks, Swin transformer consists of 4 blocks, The neck uses FPN method

결함 검출은 소규모 데이터인 것과 결함의 크기가 작다는 것을 고려하여 one-stage/two-stage 방법 및 CNN/Transformer가 적용된 backbone 네트워크에 따른 검출 알고리즘 성능을 비교했다. Backbone 네트워크에서 CNN은 ResNet을 기준으로 설정했으며, Transformer는 Swin transformer를 사용했다. 객체 검출 방법에서 one-stage는 Retinanet을 기준으로 설정했으며, two-stage는 Faster R-CNN과 Cascade R-CNN을 각각 학습하여 결과를 확인했다.

딥러닝 객체 검출 모델은 특징을 추출하는 backbone과 직접적인 분류를 수행하는 head, 두 구역을 연결해주는 neck으로 구성된다. Fig. 3과 같이 ResNet 50과 101은 5개의 블록으로 설정했으며, Fig. 4에 나타난 two-stage 모델의 backbone으로 적용되면 2번부터 5번 블록까지의 특징맵이 feature pyramid network (FPN)[23]에 적용되어 RPN과 분류 모듈인 region of interest (RoI)에 입력으로

사용된다. one-stage에서도 2번부터 5번 블록의 특징맵이 FPN에 적용되고 일정 head를 통해 bounding box의 위치와 클래스가 추정된다.

실험에 사용된 Swin transformer 모델은 Fig. 3과 같이 구성되어 있으며, 1번, 2번, 4번 블록에는 2개의 Swin transformer를 연결하고, 3번 블록에는 6(tiny)개로 구성하여 층의 깊이를 확장했다. Backbone 모델은 two-stage인 Cascade R-CNN의 경우, Faster R-CNN과 이미지 특징 추출 구역을 해당 모델과 RPN의 구성은 동일하지만, RoI가 3단계로 구성된다. backbone과 head 사이에 neck으로서 2~4번 블록이 FPN에 적용된다. Retinanet도 동일한 블록이 FPN의 특징맵으로 적용되며, 최종 추정은 일정 head를 통해 수행된다.

Fast R-CNN에서 최적 입력 이미지의 크기는 중형비와 GPU 메모리를 고려하여 짧은 축의 길이를 600픽셀로 설정했다. 이후 컴퓨터 비전 인식

분야에서는 이전 연구들과 GPU 메모리를 고려하여 최대 크기를 최대 1,333×800 범위로 제한해 모델을 학습하고 있다. 실험에서는 모델의 최적 설정값을 유지하기 위해 평균 7,149×2,237 크기를 가지는 방사선 투과 이미지를 1,333×640으로 나누어 사용한다. 128픽셀 크기로 오버랩(overlap)하여 결합이 누락되는 현상을 방지하고, 크기가 부족한 이미지는 Fig. 5와 같이 127.5픽셀 강도의 패드(pad)를 추가하여 동일한 크기로 구성했다. 이에 따라 원본 이미지는 1,338장에서 2,091장으로 증가했으며, 결합의 수가 2,751개에서 2,924개로 늘어나 데이터 증강 효과를 볼 수 있었다. 나누어진 이미지는 딥러닝 알고리즘에서 학습의 성능을 높일 수 있도록 일괄된 전처리 과정을 거친다. [24]에 따라 히스토그램 균일화, 노이즈 제거, 임계값 처리를 순차적으로 거쳐 결합의 특징은 강조하고, 배경부는 균일화하여 일차적인 분류 과정을 적용했다.

모든 모델에 대해 공통으로 입력 이미지를 1,333×640과 1,333×800 scale에 대해 다중으로 학습하는 multi-scale training을 사용했으며, 최적화 기법으로 AdamW[25], 초기 학습률 0.0001, 학습률 감소 0.0001로 설정했다. 또한, soft-NMS [26]를 통해 중복되는 후보군의 기준점을 완화했다. 비교 대상이 되는 모델은 Python 언어로 Pytorch 오픈소스 라이브러리를 활용하여 프로그래밍했다.

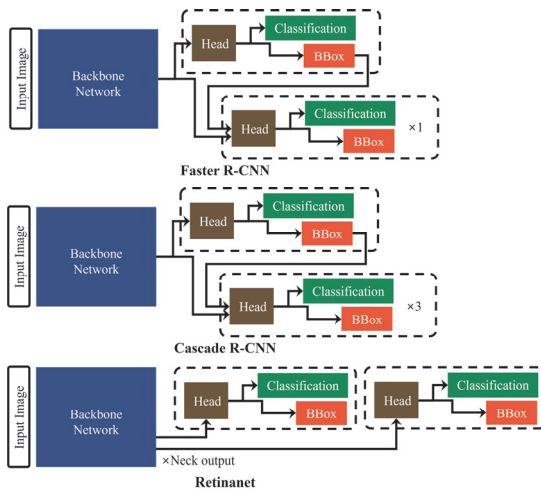


Fig. 4 Structure of three object detection models. Faster R-CNN (top) consists of one RPN head and one RoI head. Cascade R-CNN (middle) consists of one RPN head and three RoI heads. Retinanet (bottom) consists of subnets for each output of the neck

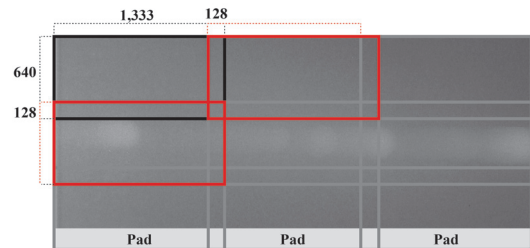


Fig. 5 Image patching strategy. Applies padding to the original image to match size during image preprocessing

4. 실험 결과

사전 학습된 가중치를 사용하는 것은 모델이 대규모의 데이터 세트에서 학습된 일반화 모델을 사용하는 것과 같으며, 새로운 데이터에 대해서도 효과적인 성능을 보장받을 수 있다. 하지만, 일반적인 대규모 데이터 세트는 3채널(red, green, blue)로 구성된 이미지다. 데이터 세트에 포함된 이미지 픽셀 강도의 평균과 표준편차를 계산하면 다음과 같다.

$$m_R m_G m_B = [0.388, 0.388, 0.388] \quad (1)$$

$$s_R s_G s_B = [91.754, 91.754, 91.754] \quad (2)$$

방사선 투과 이미지 특성상 x-ray의 배경은 0 (검은색)으로 표현되며, 용접 표면부에 대해 255 (흰색)의 값으로 나타나기 때문에 표준편차가 일반 이미지보다 크다. 또한, 1채널의 회색조(grayscale)로 인해 이전 색상 데이터에 맞추어 일반화하기 어렵다. 따라서 본 연구에서는 사전 학습된 모델이 없는 Swin transformer 입력 데이터에 대해

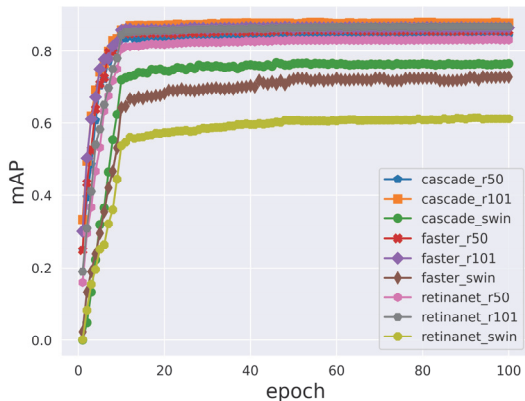


Fig. 6 mAP according to the epoch of the test data. After approximately 11 epochs, all models tend to converge

식 (1)과 식 (2)의 평균과 표준편차를 사용하여 정규화했으며, ResNet은 사전 학습 데이터에 따라 식(3)과 (4)의 값을 사용했다.

$$m_R m_G m_B = [123.675, 116.25, 103.53] \quad (3)$$

$$s_R s_G s_B = [58.395, 57.12, 57.375] \quad (4)$$

총 9개의 모델을 100 epochs로 학습했을 때, Fig. 6과 같이 대부분 11 epochs에서 최대 mAP에 수렴하는 경향이 나타난다. 공식적으로 제공되는 오픈소스에서 추천하는 학습률 조정 구간은 9 epochs와 11 epochs이며, 학습 종료 시점은 20 epochs로 설정되어있다. 이 점을 고려했을 때, 학습 과정의 경향이 일치하여 충분한 학습이 이루어진 것을 확인할 수 있다. 사전 학습된 모델이 없는 Swin transformer의 경우, 초기 mAP가 0.01 이하에서 시작하는 반면, 사전 학습 모델이 있는 ResNet 구조는 평균 0.2 이상의 초기 mAP를 보인다.

Table 2에 정리한 결과는 COCO에서 제시한

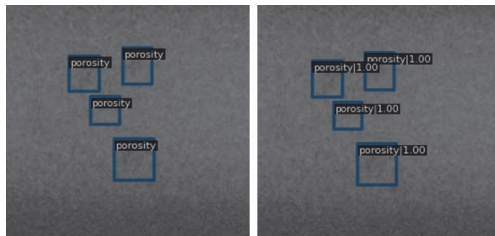
Table 2. mAP(IoU=0.50) and mAP(IoU=0.50:0.95) according to method and backbone. In the evaluation process, IoU=0.50 is set as a standard, and the average accuracy is calculated by increasing the 0.05 steps from 0.50 to 0.95 to increase the accuracy of the bounding box(IoU=0.50:0.95).

Method	Backbone	mAP _{0.50}	mAP _{0.50:0.95}
Faster R-CNN	R-50	0.856	0.654
	R-101	0.865	0.675
	Swin	0.728	0.405
Retinanet	R-50	0.831	0.605
	R-101	0.868	0.639
	Swin	0.612	0.291
Cascade R-CNN	R-50	0.842	0.674
	R-101	0.878	0.720
	Swin	0.765	0.525

기준에 따라 mAP(IoU=0.50) 와 mAP(IoU=0.50:0.95)로 나누어 성능을 확인했다.

Backbone 비교 one-stage/two-stage 방법 모두 ResNet 101을 backbone으로 구성했을 때 성능이 가장 높게 평가되었으며, ResNet 50과 Swin transformer 순으로 순위가 나열된다. Retinanet에서는 CNN에 해당하는 ResNet 101과 Transformer인 모델의 차이가 0.256으로, 가장 두드러진 차이를 보였으며, 다른 방법에 비해 mAP(IoU=0.50:0.95)가 0.291로 나타나 모델 적합도가 떨어지는 것을 확인했다.

Method 비교 mAP(IoU=0.50) 기준 Faster R-CNN과 Cascade R-CNN에서 평균 0.816, 0.828로 근소한 성능 차이를 보이나, mAP(IoU=0.50:0.95) 기준으로 평균 0.578, 0.640을 비교하면 0.062의 성능 차이를 보인다. 각 방법에서 최대 성능을 보이는 ResNet 101 기준으로 비교하면 mAP(IoU=0.50)는 0.865, 0.868, 0.878로 큰



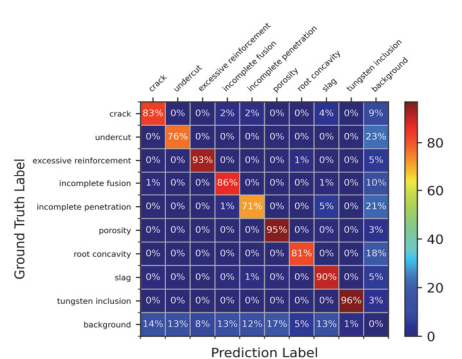
(a) Porosity



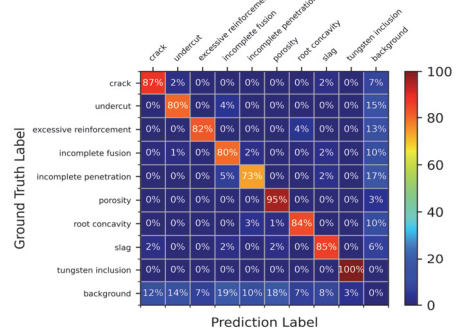
(b) Slag

Fig. 7 Weld defect detection comparison results. The top is the ground truth, and the bottom is the detection result

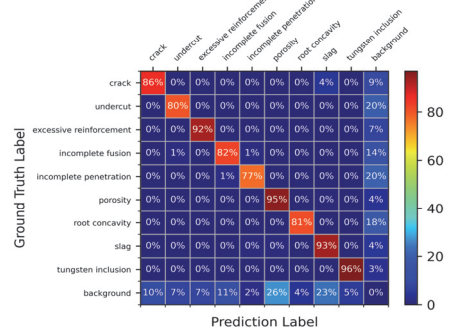
소한 성능 차이를 보이지만, mAP(IoU=0.50:0.95)는 최대 0.081의 차이로 Cascade R-CNN과 ResNet 101을 결합한 모델이 가장 높은 성능을



(a) Faster R-CNN + ResNet101



(b) Retinanet + ResNet101



(c) Cascade R-CNN + ResNet101

Fig. 8 Confusion matrix by method based on ResNet 101. The accuracy for each defect is expressed as a percentage, and the color of the cell is given according to the accuracy

보인다. Fig. 7은 해당 모델을 통해 검출한 결과다.

ResNet 101 기준으로 각 방법에 대해 용접 결함 검출 정확도를 혼동행렬로 정리하면 Fig. 8과 같이 나타난다. One-stage/two-stage 방법들 모두 결함 간의 오검출률은 5% 이하로 낮았지만, 대부분 배경부와 혼동이 일어났다. 단일 결함과 배경부를 고려하여 1종, 2종 오류를 살펴보면, tungsten inclusion이 가장 낮은 오류를 보였다. Porosity는 모든 모델에서 95%의 높은 정확도를 보였지만, 전체 1종 오류에서 큰 비중으로 나타났다. 특히 mAP 성능이 가장 높은 Cascade R-CNN에서 26%로 다른 모델보다 약 9% 높은 수치이다. 2종 오류에서 porosity는 비중이 작게 나왔지만, incomplete penetration과 undercut의 비중이 크게 나타났다. 해당 결함은 정확도에서도 최하위로, 평균 73.67%와 78.67%로 나타났다.

5. 결론

본 연구에서는 방사선 투과 이미지에서의 용접 결함을 분류하고, 용접 결함의 특성을 분석하여 이미지 전처리를 수행했다. 용접 결함을 자동 검출하기 위해 딥러닝 기반 객체 검출 알고리즘인 one-stage/two-stage 방법과 backbone을 조합하고, 각 모델의 검출 성능을 비교했다.

Backbone으로 사용된 CNN과 Transformer는 모든 방법에서 CNN의 성능이 높게 나타났다. 초기 mAP가 상대적으로 높은 것과 관련하여, 방사선 투과 이미지 데이터에서도 대규모의 데이터 세트로 학습된 일반화 모델이 효과적으로 적용된다는 것을 확인했다. One-stage인 Retinanet은 결함의 분류 정확도가 상대적으로 낮으며, two-stage인 Cascade R-CNN 구조 특성상 IoU의 수준을 높여 단계별 학습하기 때문에 0.5 기준으로

학습되는 Faster R-CNN보다 높게 나타났다. 결과적으로 모델의 검출 성능을 비교한 결과, Cascade R-CNN과 상대적으로 깊은 층인 ResNet 101 조합이 mAP 0.868로 가장 적합한 모델로 평가됐다.

모델에서 공통으로 결함 크기의 변화가 적고, 배경부 대비 색이 뚜렷하게 나타나는 결함인 tungsten inclusion의 정확도가 가장 높게 나왔다. 모델의 특징으로 비정형화된 결함을 5% 이하의 오검출률로 분류했으나, negative에 해당하는 배경부와 혼동하는 결과를 보였다. 오검출된 이미지를 살펴보면, 필름 결함이나 디지털로 전환되면서 나타나는 노이즈를 인식하거나, 용접 경계부에 생기는 음영을 인식하는 경우가 대부분 차지한다. 용접 품질 검사 특성상 2종 오류의 영향이 크게 나타나기 때문에 전반적으로 결함에 대해 정밀한 분류 방법이 필요하다. 또한, 용접 품질 검사는 결함의 유무뿐만 아니라 수치적인 치수도 고려해야 하므로 IoU의 수준을 높여 기준을 향상할 필요가 있다.

본 연구를 토대로 추후 모델의 IoU 선정 방법을 개선하고, 용접 결함 분류 정확도 향상을 위해 구별이 어려운 결함에 특화된 이미지 전처리를 도입하여 성능을 향상할 계획이다.

사 사

본 논문은 2022년도 해양수산부 및 해양수산과 학기술진흥원 연구비 지원으로 수행된 ‘지울운항 선박 기술개발사업(20200615)’의 연구결과입니다.

참고문헌

- [1] Kim, Y., Kim, J., & Kang, S. (2019). A study on welding deformation prediction for ship

- blocks using the equivalent strain method based on inherent strain. *Applied Sciences*, 9(22), 4906.
- [2] Vilar, R., Zapata, J., & Ruiz, R. (2009). An automatic system of classification of weld defects in radiographic images. *Ndt & E International*, 42(5), 467-476.
- [3] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
- [4] Datal, N. (2005). Histograms of oriented gradients for human detection. In *Proc. 2005 International Conference on Computer Vision and Pattern Recognition (Vol. 2, pp. 886-893)*. IEEE Computer Society.
- [5] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.
- [6] Sizyakin, R., Voronin, V., Gapon, N., Zelensky, A., & Pižurica, A. (2019, June). Automatic detection of welding defects using the convolutional neural network. In *Automated Visual Inspection and Machine Vision III (Vol. 11061, pp. 93-101)*. SPIE.
- [7] Tang, Y. X., Tang, Y. B., Peng, Y., Yan, K., Bagheri, M., Redd, B. A., ... & Summers, R. M. (2020). Automated abnormality classification of chest radiographs using deep convolutional neural networks. *NPJ digital medicine*, 3(1), 1-8.
- [8] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [9] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- [10] Zhu, X., Su, W., Lu, L., Li, B., Wang, X., & Dai, J. (2020). Deformable detr: Deformable transformers for end-to-end object detection. *arXiv preprint arXiv:2010.04159*.
- [11] Chen, X., Hsieh, C. J., & Gong, B. (2021). When vision transformers outperform ResNets without pre-training or strong data augmentations. *arXiv preprint arXiv:2106.01548*.
- [12] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 10012-10022).
- [13] 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).
- [14] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988).
- [15] Cai, Z., & Vasconcelos, N. (2018). Cascade r-cnn: Delving into high quality object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 6154-6162).
- [16] 김창현, 유홍연, & 홍성훈. (2007). 용접결함 패턴인식을 위한 신경망 알고리즘 적용. *한국콘텐츠학회논문지*, 7(1), 65-72.
- [17] Ahn, H. (2013). Digital Radiography. *Journal of the Korean Society for Nondestructive Testing*, 33(1), 80-95.
- [18] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In *European conference on computer vision* (pp. 740-755). Springer, Cham.
- [19] Eggert, C., Zecha, D., Brehm, S., & Lienhart, R. (2017, June). Improving small object proposals for company logo detection. In

- Proceedings of the 2017 ACM on international conference on multimedia retrieval (pp. 167-174).
- [20] Law, H., & Deng, J. (2018). Cornernet: Detecting objects as paired keypoints. In Proceedings of the European conference on computer vision (ECCV) (pp. 734-750).
- [21] Kornblith, S., Shlens, J., & Le, Q. V. (2019). Do better imagenet models transfer better?. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2661-2671).
- [22] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.
- [23] Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125).
- [24] Yun, G. H., Oh, S. J., & Shin, S. C. (2021). Image Preprocessing Method in Radiographic Inspection for Automatic Detection of Ship Welding Defects. Applied Sciences, 12(1), 123.
- [25] Loshchilov, I., & Hutter, F. (2017). Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101.
- [26] Bodla, N., Singh, B., Chellappa, R., & Davis, L. S. (2017). Soft-NMS--improving object detection with one line of code. In Proceedings of the IEEE international conference on computer vision (pp. 5561-5569).

(접수: 2022.07.25. 수정: 2022.08.04. 게재확정: 2022.08.05.)