

PrRoI Pooling을 이용한 딥러닝 기반 강판 표면 결함 탐지

정효진¹ · 이원희¹ · 허영범^{1 2†}

¹ 인하대학교 산업경영공학과

² 주식회사 와이비리서치

Deep learning based steel plate surface defect detection with Precise RoI Pooling

Hyojin Jung¹ · Wonhee Lee¹ · Youngbum Hur^{1 2†}

¹ Department of Industrial Engineering, Inha University, Incheon, South, Korea

² YB Research, Incheon, South Korea

ABSTRACT

Purpose: The purpose of this study is to attempt to improve the steel industry by suggesting ways to improve the quality of steel products.

Methods: An attempt was made to improve the detection of defect data on the surface of steel plates through data augmentation techniques and region-of-interest pooling techniques. The tools used in this study are categorized into three dimensions: object detection model, region of interest pooling, and Mixup data augmentation.

Results: The results of this study are as follows We used Mixup and Precise RoI Pooling to improve the detection performance of steel surface. We studied the effect of Mixup on the detection of steel surface defects through experiments, and found that the detection performance of certain classes is degraded when Mixup is applied. To solve this problem, we found that it is effective to apply Precise RoI pooling to improve the detection performance of the model without applying Mixup, and then we integrated Precise RoI pooling and Mixup to improve the detection performance of steel surface defects for all classes. The proposed method was found to take 0.01 seconds per sheet, which is faster than visual inspection, which takes 6 seconds per sheet.

Conclusion : Improved detection of steel plate surface defects

Key Words : Steel Plate Surface Defect, Anomaly Detection, Object Detection, Precise RoI Pooling, Mixup

† Corresponding Author(youngbum.hur@inha.ac.kr)

1. 서론

철강 제품은 제조업에서만 아니라 다양한 분야에서 주된 재료로서 사용이 되고 있는 만큼 철강 제품의 품질관리는 매우 중요하다. 다양한 철강 제품 중 판 형태의 강판 제품은 제조 공정에서 다양한 결함이 발생한다. 강판 제조 공정 중 하나인 열연 공정은 강판을 원하는 두께로 가공하는 압연 공정으로, 고온에서 압연이 이루어지는 과정이다. 이 과정에서 발생한 결함을 조기에 검출하지 못하면 이후 공정에서도 품질이 더욱 저하될 가능성이 높다. 따라서 강판 품질 관리를 위해 제조 과정에서 발생하는 결함을 신속히 발견하고, 적절히 대응하는 것이 필수적이다. 전통적으로 강판의 표면 결함 검출은 육안 검사나 간단한 기계적 검사를 통해 결함을 확인했지만, 이러한 방식은 시간과 비용의 문제가 생기고 정확도가 떨어질 수 있다(Tang et al., 2017). 이에 따라 최근에는 딥러닝 기반의 비전 검사 기술의 발전으로 인해 강판 결함을 자동으로 검출하고 분석할 수 있는 시스템이 도입되고 있다(Akhyar et al., 2023). 비전 검사는 고해상도 카메라와 컴퓨터 비전 알고리즘을 결합하여 표면 결함을 탐지하는 기술로, 결함의 위치, 크기, 유형을 정확히 식별할 수 있다. 그러나 딥러닝 모델을 활용한 결함 탐지 시스템을 구축하는 데에는 결함 데이터의 부족이라는 한계가 존재한다. 제조 공정에서 발생하는 결함들은 품질에 심각한 문제가 될수록 결함을 줄이기 위해 제조 공정을 개선시키기 때문에 결함 데이터는 상대적으로 희소하며, 특정 유형의 결함 데이터는 충분히 확보하기 어렵다. 이러한 데이터 부족으로 인한 결함 탐지 시스템 구축의 문제를 해결하지 않으면, 결함 탐지 성능이 저하될 수 있기 때문에 데이터를 증강하는 기법을 활용하여 딥러닝 모델 학습을 위한 데이터를 보강하는 것이 필요하다. 또한, 충분한 데이터가 있더라도 딥러닝 모델이 학습 과정 중에 결함에 대한 정보를 처리하는 과정 중에 모호한 처리로 정보 손실로 이어진다면, 이는 데이터 증강의 효율성 문제와 결과적으로 성능의 문제로 이어질 수 있다. 이를 위해서는 기존 모델의 구조를 결함에 대한 정보를 정밀하게 처리하여 정보 손실을 최소화할 수 있는 구조로써 개선이 필요하다.

본 연구에서는 강판 표면의 결함 이미지 데이터와 결함의 위치에 대한 정보가 담겨 있는 주석 데이터를 활용한다. 강판 표면 결함을 검출하는 객체 탐지 연구에서 이미지 데이터를 증강하는 일반적인 방법으로는 회전, 뒤집기, 자르기 등의 변형 기법이 있으며, 이는 개별 이미지에 변형을 적용하는 방식이다(Sun et al., 2022). 그러나 이러한 방법은 이미지 자체의 정보 변동성이 적어 새로운 데이터에 대한 일반화 능력이 제한될 수 있으며, 이로 인해 결함 탐지 성능이 저하될 가능성이 있다. 따라서 이를 보완하기 위해 본 연구에서는 단일 이미지 변환 방식보다 다수의 이미지를 활용하는 믹스업(Mixup) 알고리즘(Zhang et al., 2017)을 적용하여 결함 탐지 성능을 향상시키고자 한다. 이후 증강된 데이터에 대한 정보 손실을 최소화할 수 있도록 정보를 정밀하게 처리하는 PrRoI Pooling(Precise RoI Pooling)을 적용하여 결함 데이터의 정보를 모델이 효율적으로 학습할 수 있는 방법을 제안하려 한다(Jiang et al., 2018). 본 연구의 기여점은 다음과 같다.

1. 기존 강판 표면 결함 탐지 연구에서 사용되는 일반적인 데이터 증강 방식과 달리, Mixup 알고리즘을 활용하여 데이터 부족 문제를 효과적으로 보완한다.
2. 기존의 관심 영역 처리 방식에서 발생하는 정보 손실 문제를 지적하고, PrRoI Pooling(Precise RoI Pooling)을 적용하여 객체 경계를 보다 정밀하게 처리함으로써 성능 개선하는 방식을 제시한다.
3. 강판 제안한 방법론을 다양한 강판 표현 결함 데이터와 객체 탐지 모델에 적용하여, 다양한 2단계 객체 탐지 모델에서의 우수한 성능을 입증함으로써 높은 범용성을 확인한다.

본 논문의 구성은 다음과 같다 제 2장에서는 강판 표면의 결함 데이터를 이용하여 진행한 연구와 객체 탐지 분야 연구에

대해 설명한다. 제 3장에서는 본 연구에서 사용할 모델과 제안하는 강판 표면의 결함 탐지 성능 개선을 위해 적용한 데이터 증강 기법과 데이터의 정보를 정밀하게 처리하는 기법에 대해 설명한다. 제 4장에서는 본 연구에서 결과 비교를 위해 사용한 평가지표와 제안한 방법론을 적용한 실험 결과에 관해 설명한다. 마지막으로 제 5장에서는 본 연구에 대한 결론을 제시한다.

2. 관련 연구

딥러닝 모델을 활용한 결함 탐지 시스템에서는 결함을 탐지하고 분류하기 위해 객체 탐지 기법을 이용한다. 객체 탐지는 이미지 내에 있는 객체의 위치를 탐지하고 유형을 분류하는 과정으로 이루어진다. 최근 연구에서는 객체 탐지를 활용하여 강판 표면의 결함을 효과적으로 탐지하고 분류하는 방법에 대한 연구가 활발하게 진행되고 있다. Zhang et al.(2024)은 불필요한 특징을 억제하는 동시에 중요한 특징에 초점을 맞추는 잔여 공간 채널 주의 모듈(Residual Spatial-Channel (RSC) Attention Module)을 사용해 강판 표면의 결함 탐지를 개선하는 모델을 제시하였다. Shi et al.(2022)는 계층적인 트랜스포머 기반 합성곱 신경망(ConvNeXt)과 합성곱 신경망 구조에 주의 모듈을 추가한 모델을 제안하였고, 모델의 구조를 변경하는 방법을 통해 강판 표면의 객체 탐지를 수행한 실험에 대해 성능이 최대 15%가 개선됨을 확인했다. 또한, Chen et al.(2023)은 생성형 모델인 Diffusion 모델의 학습 과정을 객체 탐지에 접목시킨 DiffusionDet 모델을 제시했다. 이외에도 강판 표면 결함 탐지를 위해 다양한 딥러닝 기법이 적용되고 있다. Li et al.(2024)는 다중 스케일 특징 추출 MSFE(Multi Scale Feature Extractor) 모듈과 효율적인 특징 융합 EFF(Efficient Feature Fusion) 모듈을 적용함으로써 다양한 결함 유형과 크기를 탐지할 수 있는 방식을 제시하였다.

객체 탐지 분야에서는 모델 구조를 변경하여 결함 탐지 성능을 향상시키려는 연구가 활발히 진행되고 있을 뿐 아니라, 데이터 부족 문제를 해결하기 위해 다양한 데이터 증강 기법이 적용되고 있다. 기존에는 단순한 이미지 변환(회전, 자르기, 밝기 조정 등)을 통해 데이터셋을 확장하는 방식이 주로 사용되었다(Shorten et al., 2019). 최근에는 Zoph et al. (2020)이 자동 증강 기법을 활용하여 다양한 데이터 증강 기법을 결합하고, 학습을 통해 최적의 조합을 찾는 자동 증강기를 제안하였다. 또한 Ahn et al.(2022)은 강판 표면 결함 데이터를 이용해 기존 객체 수준 데이터 증강방식(Zhang et al., 2021)에 가중 손실 함수를 적용해 데이터 불균형한 상황에서도 효과적으로 데이터 증강을 하는 방법을 제시하였다. 이외에도 다수의 이미지를 결합하여 새로운 학습 샘플을 생성하는 방법이 활발히 연구되고 있다. 특히, 믹스업(Mixup)은 두 개 이상의 이미지를 픽셀 단위에서 선형 결합하여 새로운 이미지를 생성하는 방식으로, 기존 데이터에 존재하지 않는 형태의 데이터를 학습할 수 있도록 한다. 이를 적용하여 모델의 일반화 성능을 향상시킬 수 있음을 보였으며, 이는 강판 표면 결함 탐지에서도 데이터 다양성을 높이는 데 기여할 것으로 기대된다.

기존 객체 탐지 모델들과 최근에 제안된 Sparse R-CNN 모델은 관심 영역 RoI(Region of Interest)를 처리하는 과정에서 RoI Pooling과 RoI Align을 활용하지만(Sun et al., 2021), 두 기법은 양자화와 보간 과정에서 정보 손실이 발생하는 문제가 있다. 이 문제는 믹스업을 이용하여 데이터 증강을 진행해도 정보의 손실로 데이터 증강의 효율성이 떨어질 수 있다. 이를 해결하기 위해 보다 정밀하게 관심 영역을 처리하는 PrRoI Pooling 기법이 제안되었으며, PrRoI Pooling은 Faster R-CNN, Mask R-CNN 등 다양한 객체 탐지 모델에 적용되었으며, 기존 RoI Pooling과 RoI Align에 비해 우수한 성능을 보였다.

따라서 본 연구에서는 믹스업을 이용한 데이터 증강 기법과 PrRoI Pooling을 적용한 객체 탐지 모델(Faster R-CNN, Libra R-CNN, Cascade R-CNN) 간의 결합을 통해 강판 표면 결함 탐지 성능을 극대화하는 것을 목표로 한다. 기존 데이터 증강 방식과 PrRoI Pooling을 적용하지 않은 탐지 모델의 성능을 비교하여 제안하는 방법론의 유효성을 검증할 예정이다.

3. 이론적 배경 및 제안 방법

3.1 이론적 배경

3.1.1 Mixup

강관 표면에는 다양한 유형의 결함이 발생하지만, 각 결함 유형별 데이터가 충분하지 않아 결함 탐지 시스템을 구축하는데 어려움이 있다. 본 논문에서는 부족한 결함 데이터 문제를 해결하기 위해, 두 개의 데이터를 결합하여 새로운 데이터를 생성하는 믹스업(Mixup) 알고리즘을 활용한 데이터 증강을 시도한다. 믹스업 알고리즘은 학습 데이터 중 두 개의 데이터를 랜덤하게 뽑아 하나의 데이터 합성하는 기법이다. 합성 시에 이미지는 픽셀 단위에서 선형 결합이 이뤄지고 이미지만을 합성하는 것이 아닌 라벨 또한 같이 합성하게 된다. 믹스업 알고리즘에 대한 수식은 수식 (1), (2)와 같다. 수식 (3)의 베타 분포를 통해 랜덤하게 추출된 λ 값을 통해 두 개의 데이터를 λ 의 비율로 합성하게 된다. 예를 들어, λ 값이 0.5일 경우, 두 이미지를 각각 0.5의 비율로 선형 결합하게 된다. 이때, 첫 번째 데이터의 라벨이 (1,0)이고, 두 번째 데이터의 라벨이 (0,1)이라면, 믹스업을 적용하여 생성된 데이터의 라벨은 (0.5, 0.5)가 된다. 믹스업을 이용한 이미지의 예시는 이후 4장 실험 결과에서 제시하도록 한다.

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j \quad (1)$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j \quad (2)$$

$$\lambda \sim \text{Beta}(\alpha, \alpha) \quad (3)$$

3.1.2 객체 탐지 모델

본 논문에서는 객체 검출 모델 중 R-CNN 기반의 2단계 객체 검출 모델인 Faster R-CNN, Libra R-CNN, Cascade R-CNN을 사용한다. 2단계 객체 검출 모델은 물체가 위치할 수 있는 후보 영역을 제안하는 작업인 영역 제안(Region Proposal)과 물체의 유형을 분류하는 작업이 순차적으로 진행되어 정확도가 높다는 장점이 있다.

먼저, 영역 제안 단계에서는 물체가 위치할 수 있는 영역을 제안하는 작업으로 일반적으로 영역 제안 네트워크 RPN(Region Proposal Network)와 같은 알고리즘을 사용하여 수행된다. RPN의 경우 이미지를 인식하여 객체가 포함될 수 있는 영역을 식별하고 해당 영역 주변에 경계 상자를 생성한다. 이후 분류 및 경계 상자 회귀 단계에서는 제안된 영역 내의 객체를 분류하고 경계 상자를 좀 더 정확하게 조정한다.

3.1.2.1 Faster R-CNN

Faster R-CNN은 2단계 객체 검출 모델 중 가장 잘 알려진 모델로 기존 Fast R-CNN 모델의 학습 및 검출 속도 문제를 개선한 모델이다(Ren et al., 2015). 모델의 구조는 Figure 1와 같다.

Faster R-CNN의 구조는 크게 3가지로 나눌 수 있다. 입력 이미지의 특징맵(Feature Map)을 추출하는 백본(Backbone), 영역 제안 네트워크(RPN), 제안한 후보 영역에 대한 필터링을 통해 객체를 분류하고 경계 상자를 조정하여 결과를

도출하는 과정이 있다.

먼저, 입력 이미지는 백본 단계를 거쳐 이미지에 대한 특징을 담고 있는 특징맵이 추출된다. 추출된 특징맵이 RPN을 거치면 미리 정의된 다양한 크기의 경계 상자를 통해 객체가 있을 것 같은 후보 영역을 제안한다. 필터링 과정을 통해 제안된 후보 영역 중 중복되는 제안을 제거하여 신뢰도 높은 후보 영역만을 유지한다. 분류(Classification)와 경계 상자 회귀(Bounding Box Regression)는 완전 연결 계층(Fully Connected Layer) 거쳐 진행이 된다. 이를 위해서는 모두 같은 크기의 특징맵이 필요하다. 따라서 완전 연결 계층을 거치기 이전에 RoI Pooling을 통해 제안된 후보 영역들을 모두 같은 크기의 특징맵으로 처리한다. 이후에는 객체가 어떤 유형일지 분류하는 과정과 경계 상자 회귀 과정을 통해 경계 상자가 실제 객체에 더 잘 맞도록 조정하는 과정을 통해 최종적으로 유형, 신뢰 점수, 경계 상자 좌표가 결과로서 도출된다.

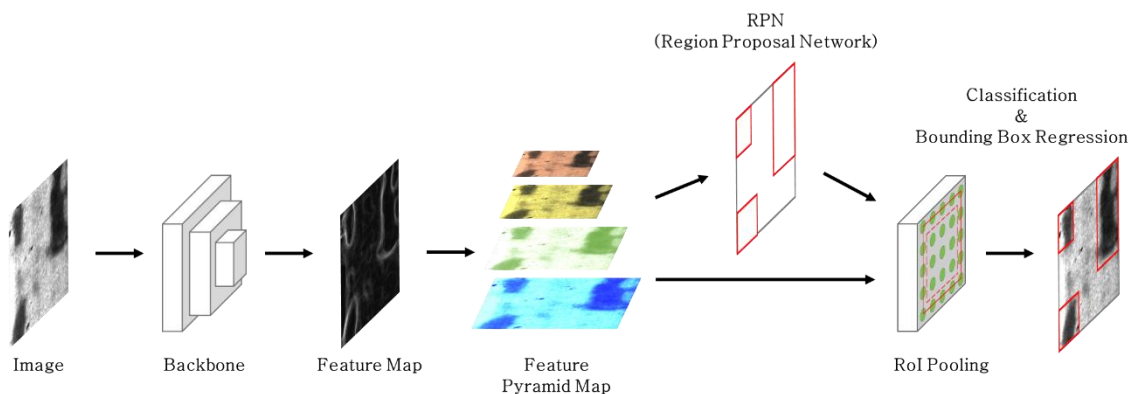


Figure 1. Object Detection Algorithm using Faster R-CNN

3.1.2.2 Libra R-CNN

Libra R-CNN은 객체 탐지 모델의 학습 과정에서 발생하는 샘플 수준, 기능 수준, 목표 수준 3가지에서의 불균형을 해결하여 기존 객체 탐지 모델을 개선한 모델로 작동방식은 동일하다(Pang et al., 2019). 모델의 구조는 Figure 2과 같다.

Libra R-CNN은 Faster R-CNN과 동일하게 입력 이미지는 백본 단계를 거쳐 이미지에 대한 특징을 담고 있는 특징맵이 추출된다. 하지만 기존 객체 탐지 모델의 백본에서 사용되는 특징 피라미드 네트워크 FPN(Feature Pyramid Network)과 다르게 Libra R-CNN의 경우 균형 잡힌 특징 피라미드 네트워크 BFPN(Balanced Feature Pyramid Network)를 이용한다. FPN에서는 입력 데이터에 대해 고수준부터 저수준의 특징맵을 추출한다. 가령, 강아지 이미지가 입력일 때 고수준에서는 강아지의 귀와 같이 큰 특징을 위주로 추출하고, 저수준에서는 강아지의 이미지가 가지고 있는 곡선과 같이 형태만을 갖고 있는 특징을 추출한다. 기존 FPN은 입력층과 가까운 고수준 특징맵에 더 큰 중요도를 부여하면서 특징 추출 과정에서 불균형이 발생할 수 있다. 이를 해결하기 위해, Figure 3과 같이 다양한 크기의 특징맵을 동일한 크기로 조정한 후 평균화하고, 다시 원래의 다양한 크기로 변환하는 과정을 적용한다. 이를 통해 학습 시에 일관되고 균형 잡힌 특징맵을 보장하여 다양한 크기의 객체에 대한 감지 성능이 향상한다.

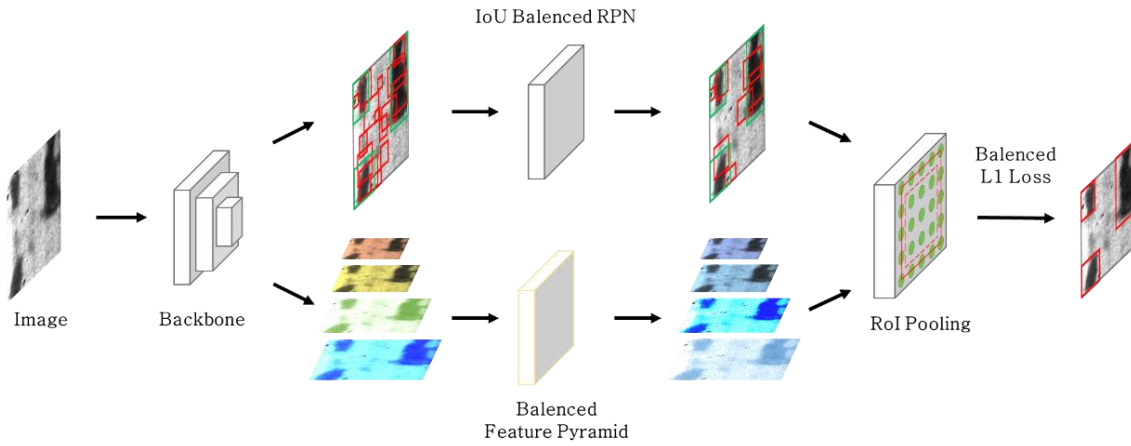


Figure 2. Object Detection Algorithm using Libra R-CNN

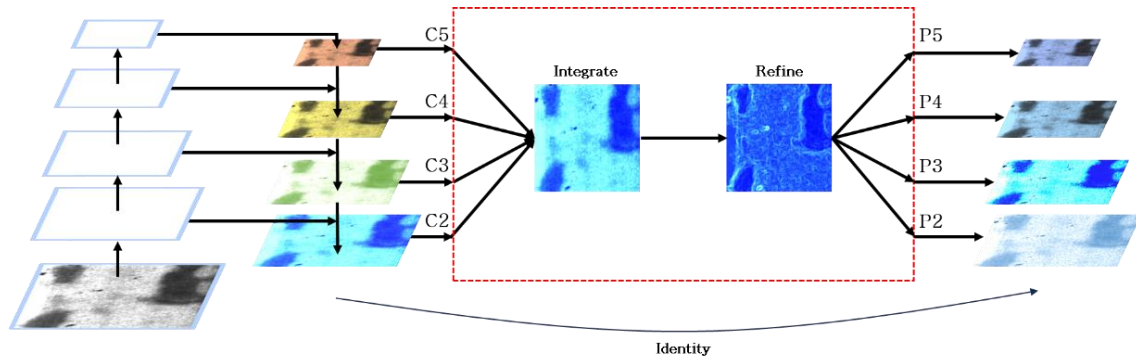


Figure 3. Pipeline and heatmap visualization of balanced feature pyramid

기존 RPN에서는 후보 영역을 제안하는 과정에서의 무작위 샘플링은 실제 영역과 후보 영역의 관련성이 높은 ‘쉬운’ 샘플을 선택하고 관련성이 낮은 ‘어려운’ 샘플은 무시하는 경향이 있다. 결론적으로 실제 영역과 후보 영역의 관련성이 낮은 어려운 데이터의 경우 탐지의 성능이 떨어진다. 이 문제를 해결하기 위해 RPN에서 후보 영역을 제안하는 과정에서의 무작위 샘플링을 IoU 균형 샘플링을 대체하여 샘플을 추출하였다. IoU(Intersection of Union)의 수식은 (4)와 같다. IoU 임계값의 구간을 나누어 임계값이 높은 ‘어려운’ 샘플과 임계값이 낮은 ‘쉬운’ 샘플 모두 동일한 수의 샘플을 추출하여 학습 안정성을 향상시켰다.

$$IoU = \frac{Overlapping\ Region(교집합\ 영역)}{Combined\ Region(합집합\ 영역)} \quad (4)$$

이후의 단계에서 남은 후보 영역들을 모두 같은 크기의 특징맵으로 처리하기 위해 RoI Pooling을 진행한다. 객체가 어떤 유형인지 분류하는 과정에서도 쉬운 샘플에 대한 과도한 의존을 방지하기 위해 어려운 샘플에 대한 가중치를 부과하는 균형 손실함수(Balanced L1 Loss)를 사용하여 일반화를 향상시킨다. 마지막으로 경계 상자 회귀 단계를 거치며 경계 상자가 실제 객체에 더 잘 맞도록 조정하는 과정을 통해 최종적으로 유형, 신뢰 점수, 경계 상자 좌표가 결과로서 도출된다.

3.1.2.3 Cascade R-CNN

Cascade R-CNN은 기존 객체 탐지 모델이 훈련을 진행할 때 고정된 IoU 임계값을 사용하여 경계 상자가 실제 객체에 더 잘 맞도록 조정하는 분류와 경계 상자 회귀를 진행하고, 이 과정이 1단계로 진행되어 성능이 저하되는 문제를 해결하기 위해 설계되었다(Cai et al., 2018). 모델의 구조는 Figure 4과 같다.

객체가 있을 것을 예상되는 후보 영역을 제안하는 RPN이 생성한 초기 영역 제안은 부정확할 수 있고, 경계 상자를 조정하는 과정이 1단계라면 잘못 예측된 경계 상자를 조정하기에 불충분한 상황이 발생할 수 있다. 이를 해결하기 위해 경계 상자를 조정하는 과정을 여러 단계에서 진행하며 후속 단계로 갈수록 점점 더 높은 임계값을 적용하여 더욱 정확하게 경계 상자를 조정한다.

기존 객체 탐지 모델은 고정된 임계값을 사용하기 때문에 특정 임계값에 과적합되어 성능이 저하될 위험이 있다. 예를 들어, 0.5와 같은 낮은 임계값을 사용할 경우, 객체가 없는 배경을 탐지하는 빈도가 증가할 수 있다. 반면, 0.7이나 0.8과 같이 높은 임계값을 적용하면, 임계값을 만족하는 샘플이 적어 학습이 어려워진다. Cascade R-CNN은 여러 단계에서 점진적으로 높은 임계값을 적용함으로써 특정 임계값에 대한 과적합을 방지하고, 보다 일반화된 모델을 학습할 수 있도록 한다.

작동 과정은 기존 객체 탐지 모델과 유사하다. 입력 이미지가 백본 단계를 거쳐 특징맵이 추출되면, RPN을 거쳐 후보 영역이 제안된다. 추출된 특징맵이 RPN을 거치면 미리 정의된 다양한 크기의 경계 상자를 통해 객체가 있을 것 같은 후보 영역을 제안한다. 다음 과정부터는 총 3번의 반복을 이룬다. 제안된 후보 영역들은 RoI Pooling을 거쳐 모두 같은 크기의 특징맵으로 처리되고 이후에는 객체가 어떤 유형인지 분류하는 과정과 경계 상자 회귀로 경계 상자가 실제 객체에 더 잘 맞도록 조정하는 과정이 진행된다. 앞선 과정을 3번 반복하고 반복을 거듭할수록 임계값이 오르게 된다. 이후 최종적으로 유형, 신뢰 점수, 경계 상자 좌표가 결과로서 도출된다.

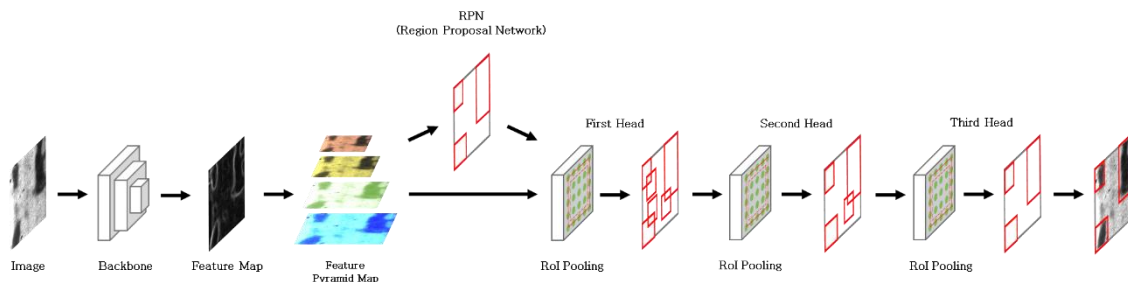


Figure 4. Object Detection Algorithm using Cascade R-CNN

3.1.3 Precise RoI Pooling

PrRoI Pooling은 기존 관심 영역 처리 기법의 정보 손실 문제를 해결하기 위해 제안된 기법이다. 기존 관심 영역 처리 기법인 RoI Pooling은 경계 상자 좌표를 양자화 하여 고정된 크기의 특징맵을 처리한다. 그 과정은 Figure 5 (a)와 같다. RoI Pooling이 적용되기 이전의 특징맵인 빨간색 점선은 RoI Pooling이 적용되면 빨간색 실선이 된다. 이 과정에서 RoI Pooling은 좌표를 반올림하여 특징맵의 크기를 고정시키는 방식으로 진행되기 때문에 좌표 정보의 손실이 발생할 수 있다. 반면, PrRoI Pooling은 정보 손실의 문제를 해결하기 위해 양자화를 제거하고, 연속적인 좌표를 기반으로 직접적으로 2차 정적분을 수행하여 보다 정밀한 특징맵을 생성한다. 그 과정은 Figure 5 (b)와 같다. 특징맵의 픽셀을 단순히 정수 좌표로 처리하지 않고, 수식 (5)를 통해 모든 연속 좌표에서 특징을 추출한다. 이후 수식 (6)을 통해 RoI 내부의 픽셀 값을 모두 더한 후 해당 영역의 크기로 나누어 평균을 계산한다.

이를 통해 모델이 정보의 손실 없이 결합의 위치와 형태를 보다 정확하게 학습할 수 있도록 한다. 예를 들어, 좌측 상단의 특징맵의 좌표가 (2.2, 3.6)이라고 할 때 RoI Pooling의 경우 (2, 4)의 좌표가 되지만 PrRoI Pooling의 경우 (2.2, 3.6)의 좌표 그대로 계산하게 된다.

$$f(x, y) = \sum_{i,j} IC(x, y, i, j) \times w_{i,j} \quad (5)$$

$$PrPool(bin, \mathcal{F}) = \frac{\int_{y_1}^{y_2} \int_{x_1}^{x_2} f(x, y) dx dy}{(x_2 - x_1) \times (y_2 - y_1)} \quad (6)$$

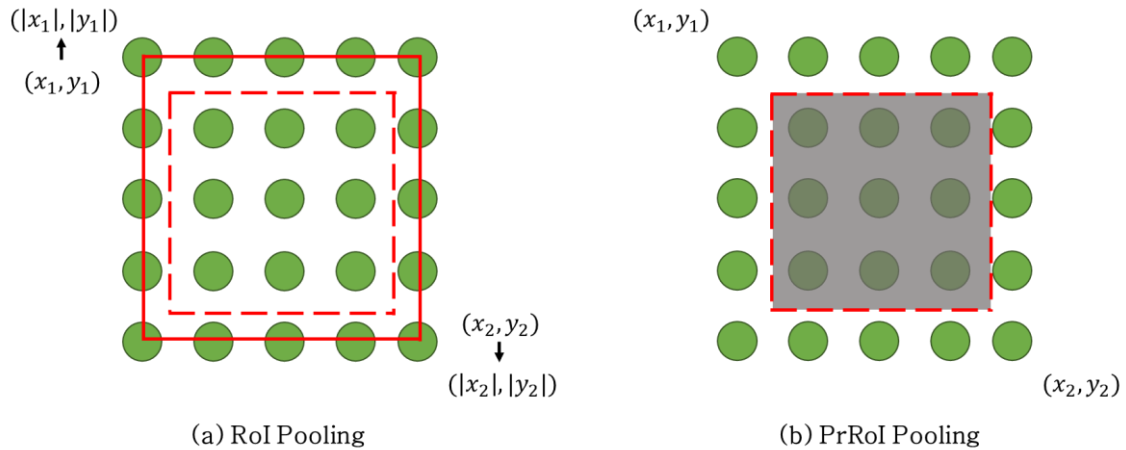


Figure 5. Illustration of RoI Pooling and Precise RoI Pooling

3.2 제안 방법론

본 논문에서 제안하는 정밀한 처리를 통한 강판 표면의 결함 탐지 방법의 구조는 Figure 6과 같다. 강판 표면의 결함 데이터가 입력으로 들어오면 믹스업을 사용하여 데이터를 증강시킨다. 이 과정에서 이미지 데이터의 경우 이미지 사이즈 또한 함께 조정한다. 이미지 사이즈를 원본 사이즈에 가깝게 하여 사이즈를 키우면 정보 손실이나 왜곡을 방지하고 탐지 성능을 향상시킬 수 있다. 믹스업을 통해 증강된 데이터와 원본 데이터는 이후 PrRoI Pooling이 적용된 모델의 입력으로써 사용된다.

2단계 객체 탐지 모델에서 PrRoI Pooling을 적용하는 과정은 다음과 같다. Faster R-CNN, Libra R-CNN, Cascade R-CNN은 모두 공통적으로 Backbone, Neck, RPN Head, RoI Head 단계를 포함하는 구조를 갖는다. 먼저, 세 가지 모델 모두 Backbone과 Neck 단계를 거쳐 특징맵을 추출하며, 이 과정에서 Libra R-CNN은 다른 두 모델과 달리 Neck 단계에서 BFPN을 활용하여 보다 균형 잡힌 특징을 추출한다. 이후, 추출된 특징맵은 RPN Head 단계를 거쳐 관심 영역 RoI(Region of Interest)을 제안받게 된다. RoI Head 단계에서는 기존의 RoI Pooling 대신 PrRoI Pooling을 적용하여 특징을 추출하며, 비교 실험을 위해 출력되는 특징맵의 크기는 RoI Pooling과 동일한 7×7 로 고정한다. 이를 통해 PrRoI Pooling이 객체 탐지 성능에 미치는 영향을 보다 정확하게 분석할 수 있다.

결과적으로, 입력 데이터는 믹스업 과정을 거쳐 증강된 후 PrRoI Pooling이 적용된 최종 모델을 통해 학습되며, 이를 통해 강판 표면 결함 탐지 성능을 향상시키는 효과를 검증한다.

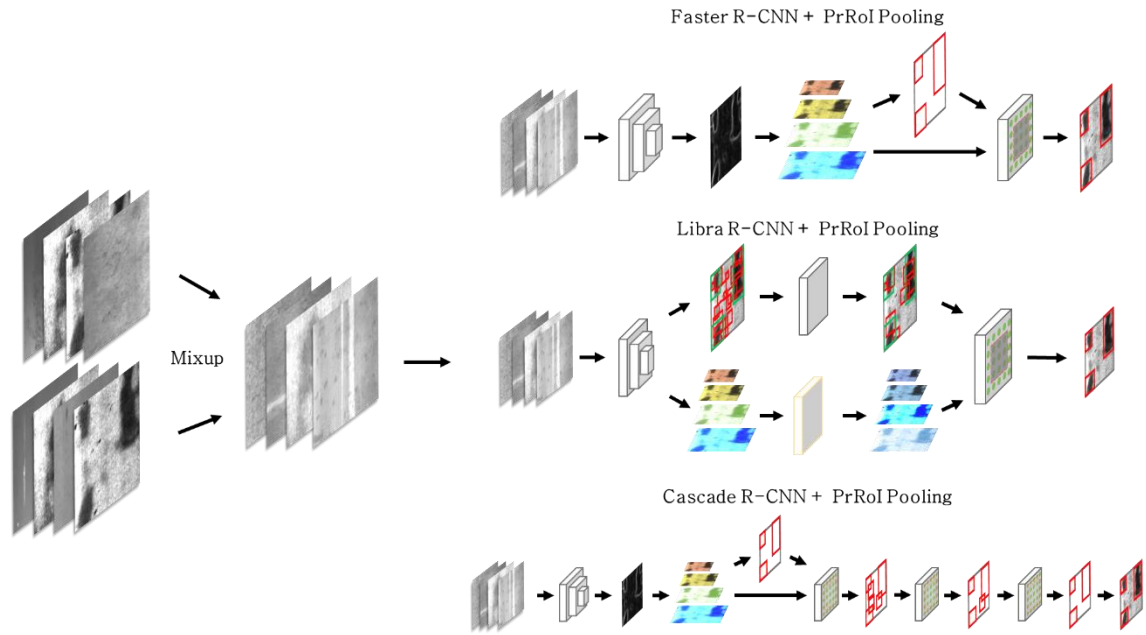


Figure 6. Illustration of suggested method

4. 실험 및 실험 결과

본 논문에서는 2단계 객체 탐지 모델에 믹스업을 이용해 증강한 데이터와 증강하지 않은 원본 데이터에 대해서도 각각 실험을 수행하고 결과를 제시한다. 이후 기존 객체 탐지 모델에 믹스업을 이용해 증강한 데이터를 넣을 경우 정보 손실로 인한 성능의 감소 유무를 확인한다. 성능 감소가 있을 경우 이에 대한 보완을 위해 PrRoI Pooling을 사용한다. 이후 이에 대한 결과를 비교위해 PrRoI Pooling을 사용 유무에 따른 실험을 각각 수행한다.

4.1 실험 세팅

본 논문에서는 Northeastern 대학에서 제공하는 NEU Surface Defect 데이터셋을 사용한다(Song et al., 2013). 데이터셋은 6개의 유형이 동일한 개수로 유형마다 300개의 JPG 형태의 이미지 파일과, XML 형태의 주석 파일 총 1800개의 이미지와 주석 파일이 있다. 유형은 압연 공정 중 강판 표면에서 발생하는 대표적인 결함 6가지이고 차례대로 Cracking(미세 균열), Inclusion(불순물), Patches(부분 얼룩), Pitted_surface(표면 함몰), Rolled-in_scale(압입된 산화철 혼합물), Scratches(긁힘)라고 한다. 이미지 파일은 200x200 사이즈의 회색조 이미지이고 주석 파일에는 이미지에 있는 결함의 위치(xmin, ymin, xmax, ymax)와 결함의 유형에 대한 내용이 포함 되어있다.

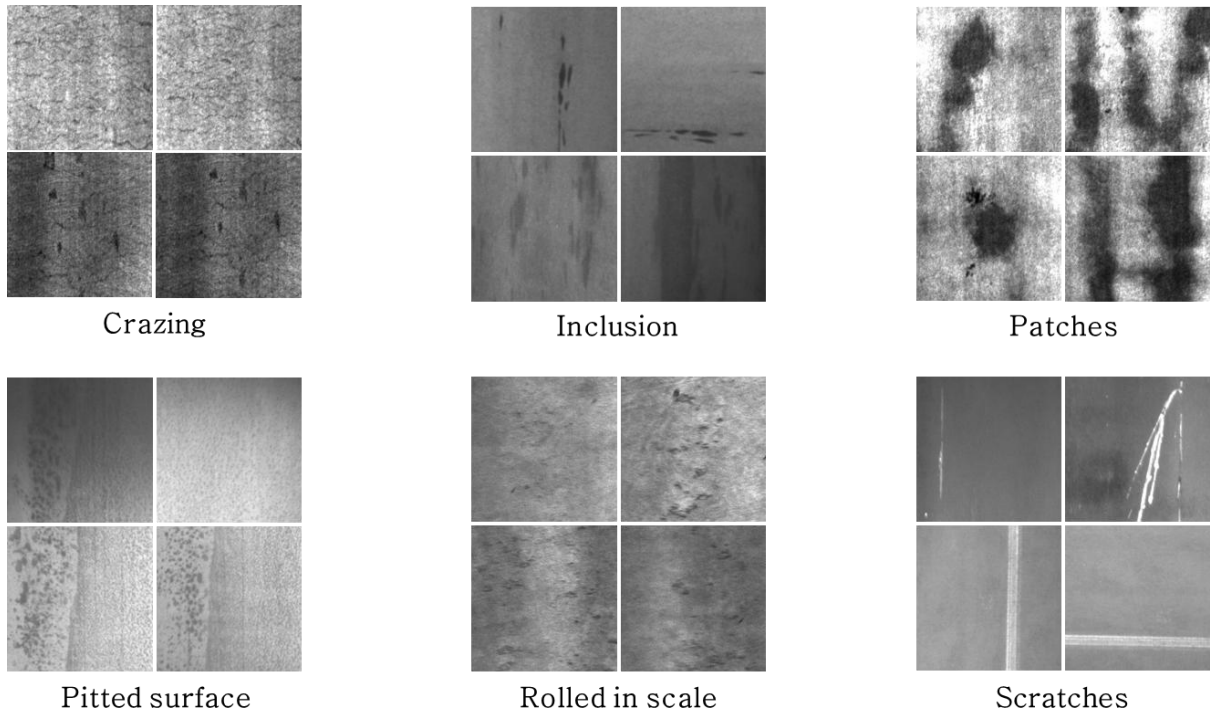


Figure 7. Samples of six kinds of typical surface defects on NEU surface defect database

데이터셋을 모델의 학습과 평가를 위해 훈련 데이터와 시험 데이터를 8:2로 나누어 학습 데이터는 유형별 240개씩 총 1440개, 시험 데이터는 유형별 60개씩 총 360개로 구성하였다. 실험에 사용하는 3개의 객체 탐지 모델(Faster R-CNN, Libra R-CNN, Cascade R-CNN)에 대해 공통적으로 설정한 세팅은 Table 1과 같다. Libra R-CNN의 경우 Faster R-CNN과 달리 Neck 단계에서 일반 FPN이 아닌 BFPN을 사용하고 이후의 RoI Head 단계에서 균형 손실함수(Balanced L1 Loss)를 사용한다. Cascade R-CNN의 경우 경계 상자를 조정하는 과정을 여러 단계로 지정할 수 있는데 본 연구에서는 모델의 성능과 검출 속도를 고려하여 3단계를 사용하였다. Cascade R-CNN에서의 IoU 임계값은 첫번째 단계 : 0.5, 두번째 단계 : 0.6, 세번째 단계 : 0.7과 같이 설정했다.

Table 1. Setting of the training parameter value

Parameters	Backbone	Learning rate	Weight decay	Momentum	Epochs	Pictures/GPU
Settings	Resnet 50	0.02	0.9	0.0001	50	2

4.2 평가지표

AP(Average Precision)는 객체 탐지 모델에 사용되는 평가지표로 모델이 객체를 얼마나 잘 탐지하고 분류하는지 정량화한다. AP는 재현율(Recall) 값의 변화에 따른 정확도(Precision)의 값을 나타낸 곡선인 PR 곡선의 면적을 통해 계산이 된다. 재현율과 정확도에 대한 식은 수식 (7), (8)과 같다. AP는 1개의 유형에 대한 평가지표이고, mAP(mean Average Precision)는 모든 유형의 AP 값에 대해 평균을 계산한 값으로 식은 수식 (9)과 같다. AP 값과 mAP 값 모두 값이 높을수록 모델이 객체를 잘 탐지하고 분류하는 것을 의미한다.

$$\text{재현율(Recall)} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (7)$$

$$\text{정확도(Precision)} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (8)$$

$$mAP = \frac{1}{n} \sum AP$$

$n : \text{the number of class}, i = 1, \dots, n$

(9)

4.3 Mixup을 이용한 객체 탐지 결과

믹스업에서 이미지에 대한 혼합 비율을 의미하는 람다 값은 0.5로 설정하고, 이미지 사이즈를 256x256으로 설정하여 모델이 결함 데이터에 대한 정보를 자세하게 학습할 수 있도록 한다. Figure 8을 통해 앞서 설정한 세팅에 대한 결과를 확인할 수 있다. 오른쪽의 믹스업 이미지는 왼쪽의 두 원본 이미지가 일정 비율로 합성되어 생성된 것이다.

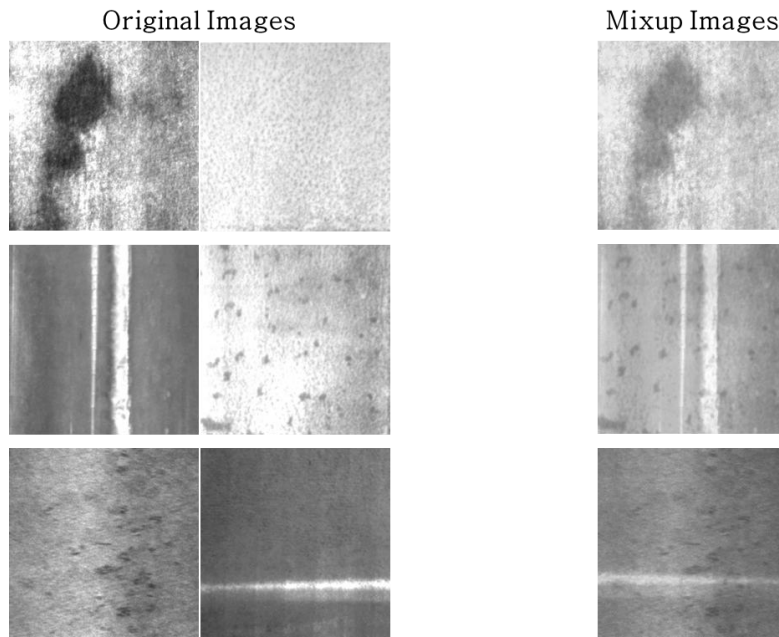


Figure 8. Examples of Original Images and Mixup Image

믹스업을 통한 데이터 증강의 유무에 따른 3개의 객체 탐지 모델에 대한 결과는 Table 2와 같다. 결과를 보았을 때, Cascade R-CNN을 제외한 Faster R-CNN과 Libra R-CNN의 mAP 값이 감소하거나 증가가 저조한 것을 확인할 수 있다. Faster R-CNN과 Libra R-CNN의 각 유형 별 AP 값을 보면 Inclusion 유형의 AP 값이 믹스업을 사용하기 이전보다 대폭 감소했기 때문으로 알 수 있다. Figure 9를 보면 Inclusion 유형은 다른 유형에 비해 작은 객체들이 많이 포함되어 있는 것을 알 수 있다. 객체의 크기가 작은 경우 몇 개의 픽셀만으로도 탐지의 유무가 결정된다. Cascade R-CNN의 경우 3개의 단계에 걸쳐 분류와 경계 상자 회귀를 진행하기 때문에 작은 객체 더라도 더 정확하게 탐지할 수 있다. 하지만 나머지 두 모델은 믹스업을 이용해 데이터 증강을 했음에도 객체의 크기가 작은 데이터에 대해서는 두 모델이 탐지를 어려워하는 것으로 보인다.

Table 2. AP (%) Comparison of Surface Defect Detection With and Without Mixup

Class	Faster R-CNN		Libra R-CNN		Cascade R-CNN	
	Without Mixup	Mixup	Without Mixup	Mixup	Without Mixup	Mixup
Crazing	39.46	40.64	41.4	46.34	37.84	47.24
Inclusion	71.76	42.76	75.3	54.96	71.34	81.64
Patches	85.16	86.28	86.74	86.28	84.36	84.94
Pitted surface	75.08	76.26	77.56	77.24	77.22	78.7
Rolled in scale	43.68	52.22	40.12	55.84	41.82	54.22
Scratches	75.08	76.36	73.68	74.68	80.98	81.12
mAP (%)	65.04	62.42	65.8	65.89	65.59	71.55

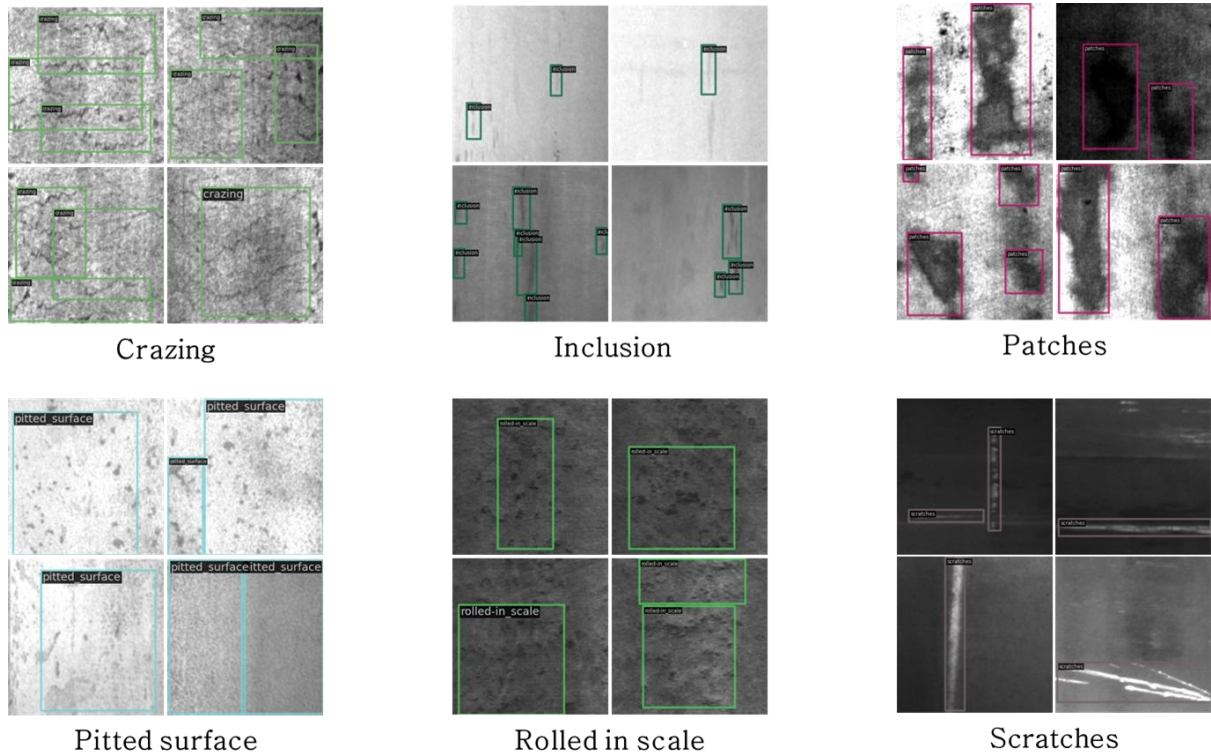


Figure 9. Test Images with Ground Truth

4.4 Precise RoI Pooling 이용한 객체 탐지 결과

4.3절에서는 관심 영역의 풀링 기법을 변경하지 않고 데이터 증강 기법만을 적용한 실험을 다루었다. Figure 10은 믹스업을 적용하지 않은 상태에서 RoI Pooling을 PrRoI Pooling으로 변경하여 객체 탐지를 수행한 결과를 보여준다. 4개의 결함 유형에 대한 결과를 살펴보면, RoI Pooling 대신 PrRoI Pooling을 사용할 경우 Ground Truth(정답)에 더 근접하게 탐지하는 경향을 확인할 수 있다. 특히, Inclusion 유형에서는 하단에 위치한 작은 객체를 보다 정확하게 탐지하였으며, Pitted surface 유형에서는 patches 유형의 결함까지 함께 탐지하는 등 전반적인 탐지 성능이 향상된 것을 확인할 수 있다.

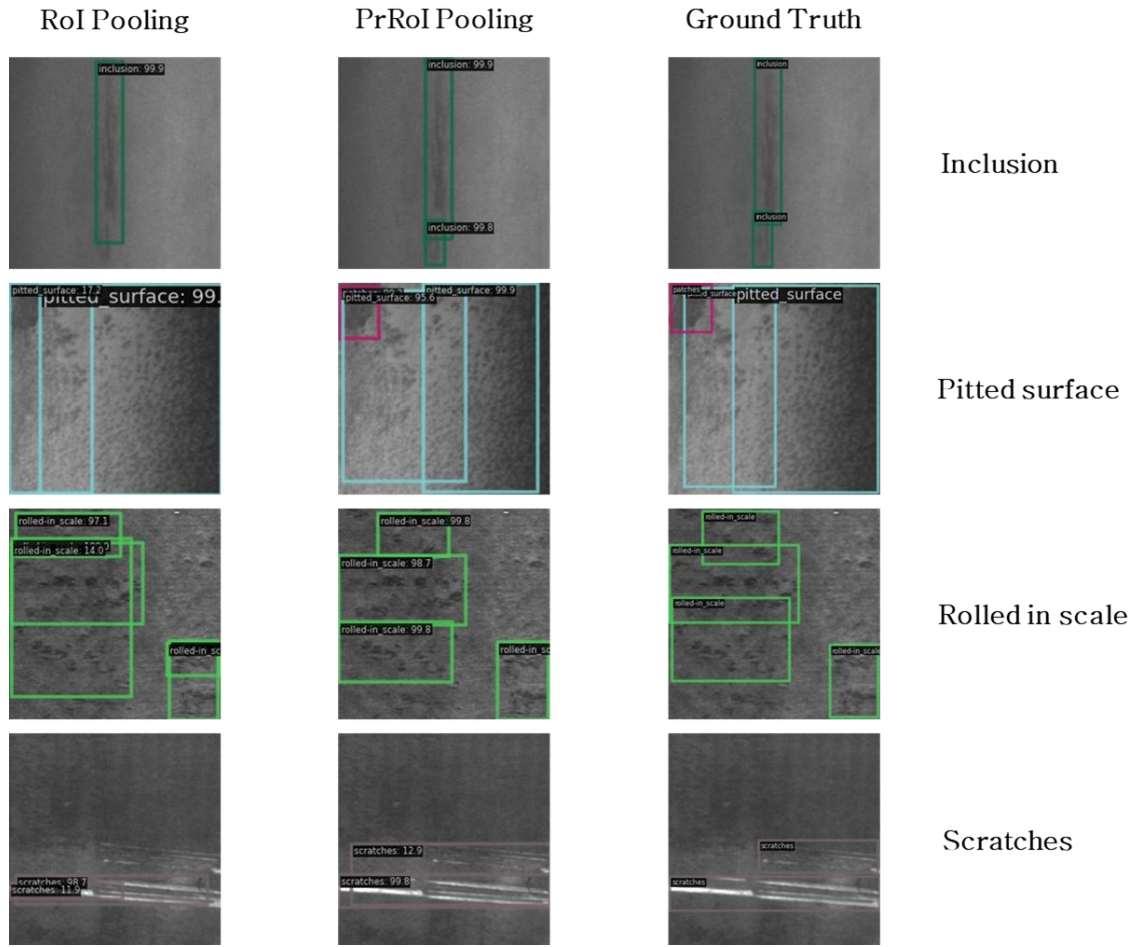


Figure 10. Ground Truth Image and Results about Faster R-CNN and Faster R-CNN using PrRoI Pooling

Table 3은 믹스업을 이용하여 데이터 증강을 수행한 후, PrRoI Pooling을 적용한 실험 결과를 나타낸다. 결과를 살펴보면, 모든 결함 유형에서 AP 값이 상승한 것을 확인할 수 있다. 특히, 믹스업 사용 이후 탐지의 어려웠던 Inclusion 유형의 경우, RoI Pooling을 PrRoI Pooling으로 변경한 후 Faster R-CNN은 약 35%, Libra R-CNN은 약 26%의 성능 향상을 보였다. 기존의 믹스업을 사용하지 않았을 때와 비교하여도, 두 모델 모두 약 6%의 성능 향상을 보였다.

Table 3. AP (%) Comparison of Mixup and Mixup + PrRoI Pooling

Class	Faster R-CNN		Libra R-CNN		Cascade R-CNN	
	Mixup	Mixup + PrRoI Pooling	Mixup	Mixup + PrRoI Pooling	Mixup	Mixup + PrRoI Pooling
Crazing	40.64	45.64	46.34	50.7	47.24	50.06
Inclusion	42.76	77.9	54.96	81.58	81.64	81.66
Patches	86.28	86.34	86.28	87.78	84.94	85.38
Pitted surface	76.26	77.68	77.24	79.3	78.7	79.5
Rolled in scale	52.22	53.44	55.84	56.12	54.22	53.7
Scratches	76.36	76.88	74.68	76.44	81.12	82.38
mAP (%)	62.42	69.65	65.89	71.99	71.55	72.11

Table 4는 제안한 방법론을 적용하기 전과 후의 실험 결과를 비교한 것이다. 믹스업을 사용하지 않고 RoI Pooling을 사용한 경우, 믹스업을 사용하고 RoI Pooling을 사용한 경우, 믹스업을 사용하고 PrRoI Pooling을 적용한 경우 결과를 3가지 모델에서 분석하였다. 분석 결과, Faster R-CNN과 Libra R-CNN은 믹스업만 사용할 경우 일부 결함 유형의 성능 감소로 인해 전체 유형의 성능인 mAP 값이 다소 감소했으나, PrRoI Pooling을 함께 적용함으로써 성능이 향상된 것을 확인할 수 있었다. Cascade R-CNN의 경우 세 번의 실험을 통해 믹스업 사용과 PrRoI Pooling 사용으로 성능이 단계적으로 올라간 것을 확인할 수 있다. 이는 믹스업과 PrRoI Pooling을 동시에 적용하면, 데이터 증강을 통해 학습 데이터의 다양성을 확보하면서도 증강된 데이터에서 발생할 수 있는 정보 손실을 최소화할 수 있기 때문이다. 또한, Table 5는 육안검사와 제안한 방법론을 적용한 모델의 이미지당 결함 검출시간이다. 장당 6초의 검출시간이 걸리는 육안검사 비해 모델의 경우 0.01초 단위의 시간이 소요됨을 확인할 수 있다(AMSquare, 2021).

Table 4. Compare mAP before and after using proposed method

mAP (%)	Using RoI Pooling without Mixup	Using RoI Pooling with Mixup	Using PrRoI Pooling with Mixup
Faster R-CNN	65.04	62.42	69.65
Libra R-CNN	65.80	65.89	71.99
Cascade R-CNN	65.59	71.55	72.11

Table 5. Detection speed about Visual Inspection and models using PrRoI Pooling with Mixup

Type	Visual Inspection	Faster R-CNN	Libra R-CNN	Cascade R-CNN
Time(sec/image)	6	0.023	0.032	0.024

5. 결 론

본 연구는 강판 표면의 결함 탐지 시스템의 구축을 위해 데이터 증강 기법과 정밀한 관심 영역 처리 기법을 결합하여 탐지 성능을 효율적으로 높일 수 있는 방법을 제안하였다. 이를 위해, 수집이 어려운 강판 표면 결함 데이터를 보완하기 위해 믹스업을 활용하여 다수의 이미지를 합성하고, 새로운 형태의 데이터를 생성하였다. 또한, 결함 영역을 보다 정밀하게 처리하여 정보 손실을 최소화하기 위해 PrRoI Pooling 기법을 적용하였다.

제안한 방법을 통해 mAP가 약 4.5%에서 최대 6.5%까지 향상되는 결과를 확인하였고, 기존 객체 탐지 모델에서 성능이 저조했던 Crazing 및 Rolled-in Scale 유형에서도 개선 효과가 뚜렷하게 나타났다. 특히, Cascade R-CNN을 적용한 경우 해당 유형에서 mAP가 약 12% 향상되는 등 제안 기법이 다양한 객체 탐지 모델에서 성능을 효과적으로 개선하는 것을 확인할 수 있었다. 또한, 제안 기법을 적용한 모델은 육안 검사 대비 약 180배 이상 빠른 속도로 결함을 검출할 수 있어 생산성 향상에 기여할 수 있으며, 작업자의 숙련도에 영향을 받지 않는 객관적인 결함 검출이 가능하다.

본 연구의 방법론은 철강 산업뿐만 아니라 결함 데이터가 부족한 다른 제조업 및 다양한 산업군에서도 활용될 가능성이 크다. 향후 연구에서는 타 산업군의 결함 데이터를 이용한 추가 실험을 통해 제안한 방법론의 일반화 성능을 검증하고, 실제 작업 환경에서 실시간 결함 탐지 시스템을 구축하여 성능을 평가할 예정이다. 또한, 경량화된 모델을 개발하여 실시간 탐지 성능을 극대화하는 방향으로 연구를 확장하고자 한다.

REFERENCES

- Ahn, Inbum, and Kim, Seoungbum. 2022. Deep Learning-Based Defects Detection of Steel Sheet Surface Using Object-level Data Augmentation. *Journal of the Korean Institute of Industrial Engineers* 48(4):327–339.
- Akhyar, F., Liu, Y., Hsu, C. Y., Shih, T. K., and Lin, C. Y. 2023. FDD: a deep learning-based steel defect detectors. *The International Journal of Advanced Manufacturing Technology*, 126(3), 1093–1107
- Cai, Z., and Vasconcelos, N. 2018. Cascade R-CNN: Delving Into High Quality Object Detection, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA. pp. 6154–6162.
- Chen, F., Sun, P., Song, Y., and Luo, P. 2023. Diffusion model for object detection. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. pp. 19830–19843.
- Jiang, B., Luo, R., Mao, J., Xiao, T., and Jiang, Y. 2018. Acquisition of localization confidence for accurate object detection. In *Proceedings of the European conference on computer vision (ECCV)*. pp. 784–799.
- Li, Z., Wei, X., Hassaballah, M., Li, Y., and Jiang, X. 2024. A deep learning model for steel surface defect detection. *Complex & Intelligent Systems* 10(1):885–897.
- Pang, J., Chen, K., Shi, J., Feng, H., Ouyang, W., and Lin, D. 2019. Libra R-CNN: Towards Balanced Learning for Object Detection, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA. pp. 821–830.
- Ren, S., He, K., Girshick, R., and Sun, J. 2015. Faster R-CNN: Towards real time object detection with region proposal networks. *Advances in neural information processing systems* 28.
- Shi, X., Zhou, S., Tai, Y., Wang, J., Wu, S., Liu, J., Xu, K., Peng, T., and Zhang, Z. 2022. An Improved Faster R-CNN for Steel Surface Defect Detection. *2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP)*. pp. 1–5.
- Shorten, C., and Khoshgoftaar, T. 2019. A survey on Image Data Augmentation for Deep Learning. *Journal of big data*, 6(1):1–48.
- Song, K., and Yan, Y. 2013. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects. *Applied Surface Science* 285:858–864.
- Sun, C., Gao, L., Li, X., & Gao, Y. 2022. A new knowledge distillation network for incremental few-shot surface defect detection. *arXiv preprint arXiv:2209.00519*.
- Sun, P., Zhang, R., Jiang, Y., Kong, T., Xu, C., Zhan, W., ... and Luo, P. 2021. Sparse r-cnn: End-to-end object detection with learnable proposals. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 14454–14463.
- Tang, B., Kong, J., and Wu, S. 2017. Review of surface defect detection based on machine vision. *Journal of Image and Graphics*, 22(12), 1640–1663.
- Zhang, H., Cisse, M., Dauphin, Y. N., and Lopez-Paz, D. 2017. mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412*.
- Zhang, J., Zhang, Y., and Xu, X. 2021. Objectaug: object-level data augmentation for semantic image segmentation. In *2021 international joint conference on neural networks (IJCNN)*. pp. 1–8
- Zhang, L., Fu, Z., Guo, H., Feng, Y., Sun, Y., and Wang, Z. 2024. TAFENet: A Two-Stage Attention-Based Feature-Enhancement Network for Strip Steel Surface Defect Detection. *Electronics*, 13(18): 3721.
- Zoph, B., Cubuk, E. D., Ghiasi, G., Lin, T. Y., Shlens, J., and Le, Q. V. 2020. Learning data augmentation strategies for object detection. In *Computer Vision-ECCV 2020: 16th European Conference, Glasgow, UK*. pp. 566–583.
- AMSquare. 2021. 인공지능 Image Segmentation 철 강판 결함 탐지 - AMSquare 기술 리뷰 1. AMSquare Blog. 2021(April). <https://zeujung.wixsite.com/mysite/post/인공지능-image-segmentation-철-강판-결함-탐지-amsquare-기술-리뷰-1-1>.