creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

공학전문석사 학위 연구보고서

딥러닝 기반 고광택 표면의 미세 결함 탐지 정확도 향상

Improvement of Micro-Defect Detection Accuracy on High-Glossy Surface Based on Deep Learning

2023년 2월

서울대학교 공학전문대학원
응용공학과 응용공학전공
고상진

딥러닝 기반 고광택 표면의 미세 결함 탐지 정확도 향상

Improvement of Micro-Defect Detection Accuracy on High-Glossy Surface Based on Deep Learning

지도교수 조남익

이 프로젝트 연구보고서를 공학전문석사 학위 연구보고서로 제출함

2023년 2월

서울대학교 공학전문대학원

응용공학과 응용공학전공

고상진

고상진의 공학전문석사 학위 연구보고서를 인준함

2023년 2월

위원장 곽 우 영 (인)

위 원 조 남 익 (인)

위 원 한 훈 (인)

초 록

기술의 발전과 최근 스마트팩토리의 구축과 함께 제조 산업 전반에 자동화가 이루어졌다. 기업은 자동화를 통해 생산성 향상, 인건비 감소의 효과를 얻었으며 딥러닝 기술의 발전으로 그 분야가 더욱 확대되어 가고 있다. 특히 합성곱 신경망을 이용한 영상 인식 분야의 딥러닝 기술의 발전에 따라 딥러닝 기반의 다양한 검사 방법이 적용되고 있는 추세이다. 하지만 여전히 제품 형상, 표면과 결함 상태, 자동 검사 신뢰성 미확보 등의 이유로 여전히 전문 인력의 육안 검사가 시행되고 있는 분야도 남아있다. 딥러닝을 통한 자동 검사를 하기 위해서는 측정 가능한 이미지를 얻을 수 있는 비전 시스템이 구축되어 있어야 하며 원하는 성능을 낼 수 있도록 학습에 충분한 양의 데이터가 필요하다.

본 연구 보고서에서는 고광택 표면의 제품 상에 미세하며 비정형 결함에 대한 이미지 취득 방법, 부족한 데이터 문제에 대한 해결 방법과 딥러닝 기반 검출 정확도 향상 방안에 대해 제안한다. 이미지 취득은 가변 조명과 표면 기울기를 측정하여 고광택 표면의 빛 반사에 의한 문제를 해결하였으며 데이터 부족에 대한 문제는 포아송 이미지 편집과 데이터 증량(Data augmentation)을 통해 학습에 충분한 데이터셋을 확보하였다. 또한 산업 현장에 적용할 수 있는 성능 신뢰성 확보를 위하여 정확도(Accuracy), 정밀도(Precision), 재현율(Recall)의 수치를 최대화 하는 것을 목적으로 하였다. 목적의 달성을 위하여 전이 학습(Transfer learning)을 적용하여 데이터 부족의 보완, 정확도 향상을 이루었다. 또한 원본 이미지 대비 결함의 크기가 작아 특징 추출 및 인식이 저하되는 것을 방지하기 위하여 이미지를 학습에 유효한 크기로 분리하여 학습하였으며 사전 학

습된 모델 별로 비교하여 가장 성능이 좋은 모델을 선정하였으며 DOE를 통한 매개변수 변경으로 최적 조건을 도출하여 정확도 99.3% 달성을 확인하였다.

주요어 : 조명, 비전센서, 데이터 증량, 딥러닝, 합성곱 신경망, 전이학습
학번 : 2021-20491

목 차

I. 서론 **1**

1.1 연구 배경 1

1.2 연구 내용 3

1.3 연구 보고서 구성 6

II. 배경 지식 및 관련 연구 **7**

2.1 표면 결함 7

2.2 머신 비전 11

2.2.1 조명 12

2.3 딥러닝과 합성곱 신경망 15

2.3.1 딥러닝 15

2.3.2 합성곱 신경망 19

2.4 합성곱 신경망 아키텍처 22

2.4.1 VGGNet 22

2.4.2 ResNet 23

2.4.3 Inception(GoogLeNet) 25

2.4.4 Xception 27

2.4.5 EfficientNet 28

2.5 전이학습 30

2.6 포아송 이미지편집 32

III. 문제 정의 및 해결 방안 **34**

3.1 문제 설명 34
3.2 해결 방안 36

IV. 실험 및 결과 39
4.1 이미지 취득 39
4.2 데이터셋 구성 및 전처리 46
4.3 이미지 합성 및 데이터 증량 47
4.4 실험 환경 설정 52
4.5 실험 결과 및 평가 53
4.5.1 아키텍처 별 전이학습 성능 비교 53
4.5.2 학습률, Batch Size 별 성능 분석 - Front 58
4.5.3 학습률, Batch Size 별 성능 분석 – Back 64
4.5.4 이미지 분할 학습을 통한 성능 향상 검증 68

V. 결론 73

참고 문헌 76

Abstract 81

그 림 목 차

그림 1. 연구 대상인 스테인리스 소재 제품 3
그림 2. 프레스 가공 종류 7
그림 3. 금속 판재 프레스 성형 시 표면 결함 유형 8
그림 4. 박판 소재의 보호비닐 부착 9
그림 5. 프레스 금형을 통한 드로잉 성형 공정 9
그림 6. 금형 내 이물에 의한 표면 찍힘 발생 과정 10
그림 7. 머신 비전 시스템 구성 12
그림 8. 후면광과 확산광의 조명 조사 방식 13
그림 9. 투명유리와 거울면에서의 빛의 반응 13
그림 10. 산란광 14
그림 11. 조명의 위치에 따른 암시야와 명시야 15
그림 12. 인공지능, 기계학습, 딥러닝의 벤다이어그램 16
그림 13. 인공신경망 17
그림 14. 합성곱신경망의 구조 20
그림 15. 합성곱신경망의 특징맵 추출 과정 21
그림 16. 풀링의 종류 별 예시 21
그림 17. VGGNet 구조 22
그림 18. Residual Block 24
그림 19. Inception Module 25
그림 20. Inception V2 26
그림 21. 단순화 한 구조의 Inception Module 28
그림 22. Xception 28

그림 23. Scaling Up 세가지 방법 29
그림 24. 데이터 규모에 따른 전이학습 방법 및 구조 31
그림 25. 포아송 이미지 편집 기법 33
그림 26. 반사의 종류 37
그림 27. 반사의 형태 및 방향과 전체 반사광 39
그림 28. 비전 센서 시스템 구성 41
그림 29. 비전 센서 내부 단면도 및 LED 배열 42
그림 30. 비전센서로부터 얻은 7가지 데이터 42
그림 31. 실험 조건 세팅 43
그림 32. Front, Back 표면 찍힘 검출 테스트 44
그림 33. 조명의 종류 별 취득 이미지 45
그림 34. 비전 센서 현장 설치 45
그림 35. 이미지 마스킹 전 후 비교 46
그림 36. 엣지 검출을 통한 이미지 전처리 47
그림 37. 초기 실험 합성곱 신경망 구조 47
그림 38. 초기 실험 합성곱 신경망 성능 그래프 48
그림 39. 포아송 이미지 편집을 통한 합성 과정 49
그림 40. 회전과 대칭을 통한 데이터 증량 51
그림 41. 사전훈련모델 결정맵 54
그림 42. Back의 Inception V3와 Xception 전이학습 성능 그래프 56
그림 43. Back 영역의 실험 설정 변경 후 전이학습 성능 그래프 58
그림 44. Front Inception V3, Xception의 과적합 조건 63
그림 45. Front Inception V3, Xception의 정상 성능 조건 63
그림 46. Back Inception V3, Xception의 정상 성능 조건 65
그림 47. Front 불량 이미지 분할 68

그림 48. Front 불필요 영역 삭제 69
그림 49. Front 이미지 분할 Inception V3, Xception의 성능 . . . 70

표 목 차

표 1. 혼동 행렬(Confusion matrix) 36
표 2. 데이터 증량 전,후 수량 비교 51
표 3. 실험 환경 설정 52
표 4. 전이학습을 위한 사전학습모델 비교표 54
표 5. 전이학습 조건 요약 55
표 6. 전이학습 정확도 비교 56
표 7. 학습률과 Batch Size 변경 후 전이학습 정확도 비교 . . . 57
표 8. 학습률, Batch Size, Epoch 조건 별 실험 계획 59
표 9. Front 실험 조건 별 Inception V3 전이학습 성능 61
표 10. Front 실험 조건 별 Xception 전이학습 성능 62
표 11. Back 실험 조건 별 Inception V3 전이학습 성능 66
표 12. Back 실험 조건 별 Xception 전이학습 성능 67
표 13. Front 이미지 분할 실험 조건 별 Inception V3 전이학습 성능 71
표 14. Front 이미지 분할 실험 조건 별 Xception 전이학습 성능 72

제 1 장

서론

1.1 연구 배경

산업이 발전함에 따라 생산하는 제품, 형태, 소재가 다양해지고 있다. 이에 따라 종래의 소품종 대량생산에서 소비자 각각의 취향에 맞춘 다품종 소량생산의 형태를 띄고 있으며 이러한 추세에 발맞추어 대량생산의 토대가 되는 금형과 자동화 기술도 함께 발전해왔다. 금형을 이용한 생산은 그 대상 소재와 방식에 따라 사출 금형, 프레스 금형, 세라믹 금형 등으로 분류되며 생산 공정은 크게 재료의 투입, 생산, 검사로 이루어지며 대부분의 공정은 자동화 시스템이 구축 되어 있다. 하지만 재료와 제품 특성에 따라 자동화가 되지 않은 경우가 있으며 이 중 검사 공정은 여전히 사람의 육안 검사가 수행되는 제품이 있다. 자동화가 어려운 이유로 대형 가전과 같이 크기가 큰 제품의 경우 검사 시스템을 구축하기 위한 비용과 현장의 조명 혹은 공간 부족 등의 환경적 문제, 비정형 형태와 측정이 불가능한 불량 유형 등 다양한 이유로 육안검사를 시행하고 있다. 사람이 육안으로 검사를 할 경우 기계가 놓칠 수 있는 미세한 결함이나 빛 반사에 의해 놓칠 수 있는 결함도 검출할 수 있는 반면 인력의 투입, 생산성 저하의 단점이 있다. 따라서 기업은 계속되는 인건비의 상승과 원가를 줄이기 위해 자동화를 위한 노력을 끊임없이 하고 있으며 기술이 발전함에 따라 기존의 육안 검사에서 기계에 의한 검사가 가능해지고 인건비 절감, 생산 시간 확대, 생산성 향상의 효과를 얻을 수 있게 된다. 자동화 검사가 가능

할 수 있도록 하는 데에 다양한 기술이 발전, 적용되어 왔으며 이 중 비전 센서와 딥러닝 기술의 발전이 큰 비중을 차지하고 있다고 할 수 있다. 인력이 아닌 기계를 통한 자동화 검사를 크게 인식과 판정으로 나누었을 때 인식에 비전 센서, 합불 판정을 위하여 딥러닝 기술이 적용될 수 있다.

제품 표면 결함의 경우 머신 비전 카메라를 사용한 비전 센서를 많이 사용하고 있다. 머신 비전 카메라는 CCD, CMOS 등의 센서로 들어온 대상물의 반사되는 빛의 양과 종류를 분석하여 원하는 정보를 얻는다. 비전 센서를 사용하여 데이터를 얻을 시 비접촉, 자동화 검사가 가능하며 최근에는 광학 기술의 발전에 따라 고속, 고해상도의 데이터를 얻는 것이 가능해졌다. 기존의 머신 비전을 이용한 검사는 카메라 혹은 비전 센서로부터 얻은 정보를 룰베이스(Rule-based)로 검사를 하는 것이 보통이었다. 통상 정형적이며 측정이 가능한 불량의 경우 룰베이스를 통한 머신 비전 방식을 통해 자동화 검사가 가능하다. 하지만 비정형적이며 미세한 결함의 경우는 룰베이스의 머신비전으로는 감지하기가 어려우며 최근에는 딥러닝을 통한 불량 검출을 시도하고 있는 추세이다. 딥러닝은 인공신경망을 기반으로 여러 개의 레이어를 사용하여 이미지 데이터의 특징을 추출하고 이를 학습하여 새로운 데이터가 입력 시 학습된 정보를 바탕으로 추론하는 기술이다. 딥러닝 기술의 발전은 이미지 처리 분야에서 비약적인 발전을 이루었으며 특히 CNN 모델의 발전으로 인하여 이미지 데이터는 경우에 따라 사람이 인지하는 수준보다 더 높은 정확도를 보인다. 이와 같은 기술의 발전으로 자동화가 이루어지고는 있지만 제품의 특징, 요구되는 수준, 현장의 환경 등 여러 제약사항들이 있으며 여전히 육안 검사를 시행하는 경우가 있다. 특히 표면 광택도가 높으며 불량의 형태가 정형적이지 않은 부품 검사가 대표적인 예이다. 비전센서와 딥러닝을 사용하여 검사를 하기 위해서는 제품 표면 이미지 취득, 학습을 위한 충분한 데이

터가 필요한데 광택이 높을 경우 조명과 채광 등에 의한 빛 반사가 심해 원하는 표면 이미지를 얻기 힘들며 미세한 결함의 경우 빛 반사에 의해 불량이 보이지 않게 된다. 또한 보통 양품보다는 불량의 수가 현저히 적어 데이터 부족이 발생할 수 밖에 없으며 이에 따라 충분한 학습이 이루어지지 못해 높은 성능을 내기 어렵다.

이러한 문제들을 해결하기 위해 본 연구에서는 빛 반사가 심한 고광택 제품의 표면 결함 이미지 취득 방안, 부족한 데이터의 증량 방법, 딥러닝을 통한 양품과 차이가 미세한 결함을 높은 정확도로 분류할 수 있는 모델 제안을 목적으로 한다.

1.2 연구 내용

본 연구의 목표는 고광택 제품의 비정형 미세 결함의 최적 검출 방법을 제시함에 있어 빛 반사가 심한 제품에 대해 비전 센서를 통한 이미지 취득 방법, 데이터 부족 문제 해결, 효율적인 특성 추출 방법 및 결함 검출 정확도 향상 방법을 제시하는 것이다. 연구에 적용한 제품으로는 [그림 1]과 같은 가전의 스테인리스 부품이며 Front, Back 으로 나뉘어져 있다.

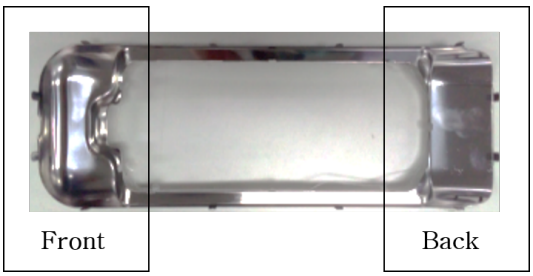


그림 1: 연구 대상인 스테인리스 소재 제품

불량의 형태는 찍힘, 긁힘 등 다양한 형태로 나타나나 이 중 한 개라도 있을 경우 불량으로 처리되기 때문에 결함의 종류를 구분하지 않고 전체적으로 불량으로 판단한다. 해당 제품은 광택도가 높고 결함의 형태, 위치가 정해져 있지 않아 머신 비전으로 검사가 불가하여 전문가의 육안 검사가 이루어 지고 있다. 이에 본 연구를 통해 자동화 검사가 가능한 방법을 제안하고자 하였으며 이에 따라 다음 네 가지에 대한 연구를 진행하였다.

첫 번째로 고광택 제품의 미세 결함 이미지 취득에 대한 연구로 비전 센서를 통한 이미지 수집 방안이다. 비전 센서를 통해 들어오는 정보는 물체로부터 반사된 빛을 정보로 받는데 표면의 상태, 반사의 종류에 따라 복합적인 형태를 보이며 통상적인 비전 센서의 경우 전체적인 반사율에 대한 정보를 얻게 된다. 하지만 고광택 표면을 가진 제품은 빛이 과도하게 반사되어 조명의 위치, 제품의 형상에 따라 표면의 정보를 제대로 얻지 못하게 되는 문제가 발생한다. 본 연구에서는 이를 해결하기 위한 방법으로 전체적인 반사율이 아닌 표면 기울기, 표면 거칠기, 반사율 등의 정보를 개별적으로 분리하여 원하는 표면 상태만 이미지로 얻을 수 있는 비전 센서를 적용하였다.

두번째로는 산업 현장에서 흔히 발생하는 결함 데이터 부족에 따른 학습과 검출 성능 저하의 대책이다. 비전 센서로 원하는 이미지를 얻을 수 있다고 하여도 불량이 과도하게 발생하지 않는 이상 결함 데이터 수는 부족할 수 밖에 없어 학습에 충분한 데이터를 얻기는 어렵다. 이러한 데이터 부족을 해결하기 위하여 데이터 증량을 진행하였다. [그림 1]의 Back의 경우 불량의 빈도가 Front 대비 높지만 양쪽 모두 절대적인 불량 데이터의 수가 부족하여 합성, 회전, 반전을 통해 증량하였다.

세번째로는 불량의 크기가 전체 이미지 대비 미세한 크기이며 비정

형적이다. 불량 데이터가 충분히 많을 경우 특성 추출이 비교적 용이할 수는 있으나 현실적으로는 데이터 증량을 하여도 충분히 많은 데이터를 얻기에는 한계가 있다. 이에 크기가 큰 이미지 내 미세한 결함을 탐지하는 방식이 아닌 이미지 분할을 통해 불량의 영역이 전체 이미지에서 차지하는 비율이 커질 수 있도록 하여 특성 추출이 용이하도록 하였다.

네번째, 전이 학습을 통한 양품과 불량의 분류 정확도를 향상하였다. 딥러닝 기술, 특히 합성곱 신경망 기술의 발전으로 이미지의 분류에 있어 높은 성능을 낼 수 있게 되었다. 충분한 학습 데이터와 적절한 모델을 적용하면 인간의 분류 능력보다 더 나은 성능을 보이기도 한다. 합성곱 신경망을 통해 깊은 은닉층과 충분한 데이터를 바탕으로 이미지가 가진 특성을 추출하여 학습하면 분류 정확도를 높일 수 있으나 산업 현장에서 현실적으로 많은 양의 데이터를 얻기란 어려우며 깊은 레이어를 학습 할 경우 학습 시간이 매우 오래 걸리게 된다. 이를 해결하기 위한 방법으로 본 연구에서는 사전 학습된 딥러닝 모델의 특성치를 적용하는 전이 학습을 적용하였으며 이를 통해 불량과 양품을 높은 정확도로 구별할 수 있는 것을 확인하였다.

본 연구에서는 고광택 표면의 이미지 데이터 취득이 가능하도록 하는 비전 센서, 부족한 데이터의 증량과 특성 추출 효율성 향상, 양품과 불량의 분류 정확도를 높일 수 있는 학습 방법에 대해 검증하였으며 이를 통해 고광택 스테인리스 제품의 미세 결함에 대한 검사 자동화를 위한 가능성을 확인하였다.

1.3 연구 보고서 구성

본 연구 보고서의 구성은 다음과 같다.

제 2장에서는 표면 결함, 표면 반사와 딥러닝 등 본 연구를 위한 배경 지식에 대해 소개하고 제 3장은 문제와 해결 방안에 대한 내용이다. 제 4장에서는 실제 실험을 위한 데이터셋과 실험 설정, 그리고 실험 결과에 대해 기술하였고 마지막 제 5장에서는 결론 및 향후 적용 방안에 대해 정리하였다.

제 2 장

배경 지식 및 관련 연구

2.1 표면 결함

프레스 성형은 금속 판재 제품의 가공 방법 중 가장 많이 쓰이는 공법 중 하나이다. 프레스 성형이란 프레스 기계를 사용하여 소재를 금형 안에 투입시켜 압력을 가한 후 원하는 형상을 얻는 성형 방법이며 이때의 금형을 프레스 금형이라 한다. 프레스 성형 제품의 경우 1개의 공정이 아닌 다수의 공정으로 제품 성형이 완료되는 것이 일반적이며 성형법으로는 [그림 2]와 같이 크게 전단, 굽힘, 드로잉 성형으로 나뉘어진다.

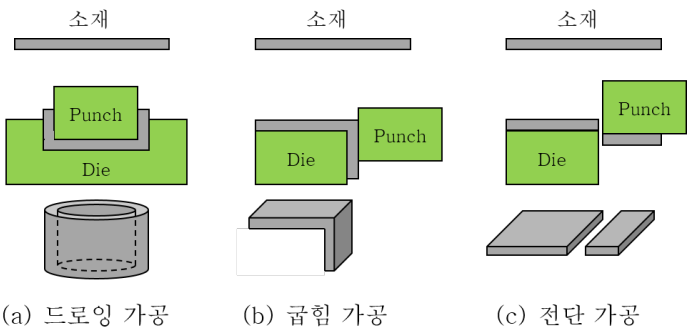


그림 2: 프레스 가공 종류

프레스 성형 시 주름, 파단, 찍힘, 스크래치 등 다양한 외관 불량이 나타나는데 외장품으로 쓰이는 제품일수록 양품 조건이 까다롭다. 주름과 파단의 경우는 발생 공정이 어느 정도는 특정되어 있으며 눈으로 식별이

가능하기 때문에 검사가 까다롭지 않은 반면 찍힘, 스크래치와 같은 미세한 결함은 불량 발생 위치도 특정하기 어려우며 검출하기도 쉽지 않다. 특히 제품 크기 대비 결함 크기가 작고 형상이 복잡한 경우 비전검사로는 찾기 어려워 전문인력이 육안검사를 해야 하는 실정이다.

표면 결함은 [그림 3]과 같은 형태로 주로 나타나며 그 발생 원인은 다양하지만 주로 이물에 의한 결함인 경우가 많다. 외장품에 쓰이는 소재 중 이송 및 취급 시 발생되는 찍힘과 스크래치를 방지하기 위해 [그림 4]와 같이 보호 비닐을 입히는 경우가 있다. 보호 비닐을 사용하면 외관의 표면을 보호하는 역할도 하면서 금형과 소재 간의 마찰계수를 낮춰주는 효과가 있어 드로잉 성형에 유리하다. 하지만 역설적으로 이 보호 비닐에 의한 찍힘이 많이 발생하게 된다.

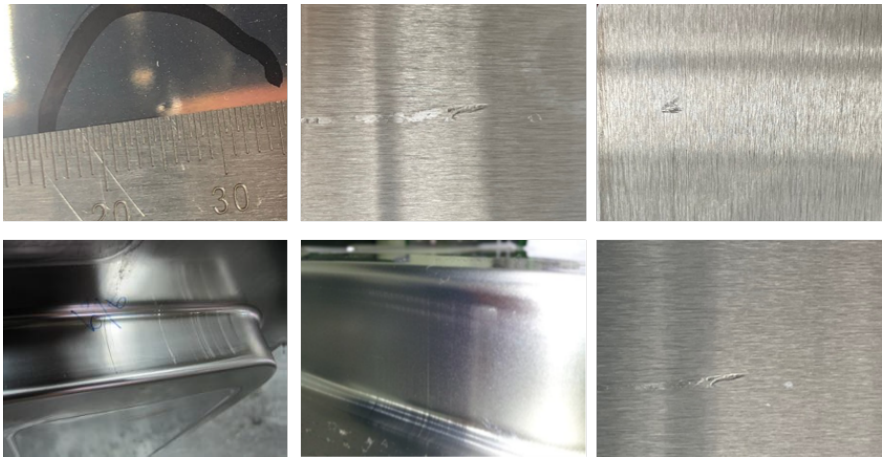


그림 3: 금속 판재 프레스 성형 시 표면 결함 유형

앞서 언급한 바와 같이 프레스 성형에 의한 제품은 프레스 금형의 가압에 의해 성형 되며 드로잉 성형과 벤딩 성형 시 [그림 5]와 같이 소재가 펀치에 밀려 다이 속으로 들어가게 되면서 금형과 소재 간 마찰과 소재의

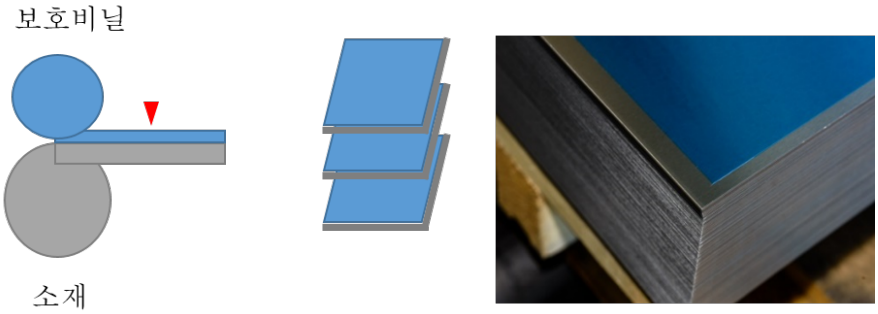


그림 4: 박판 소재의 보호비닐 부착

연신이 일어나게 된다. 금형과 소재는 금속이지만 보호 비닐은 PVC, PE 등의 폴리머로 통상 30~80㎛ 의 필름형 태로 소재에 접착 되어져 있다. 보호 비닐의 경우 연신율 자체는 400~500% 수준으로 매우 높지만 충격과 마찰에 취약하여 성형 시 보호 비닐이 찢어지는 경우가 많이 발생한다. 찢어지면서 발생되는 보호 비닐 찌꺼기가 금형 내에 남아 이 후 투입되는 소재에 찍힘을 유발하거나 손상 된 보호 비닐이 붙은 채 다음 공정으로 넘어간 소재에 의해 이 후 공정의 금형에도 이물 형태로 남게 되어 찍힘 불량의 원인이 된다.

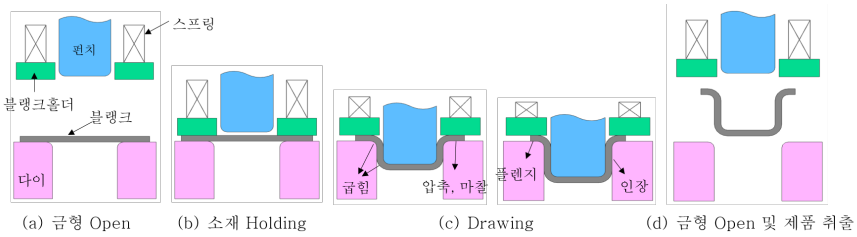


그림 5: 프레스 금형을 통한 드로잉 성형 공정

VCM(Vinyl Coated Metal) 이라 불리는 PET 층을 외관으로 사용하는 소재의 경우에도 드로잉과 벤딩 성형 시 PET 층이 박리 및 손상되는 불량이 많이 발생한다. 이에 성형 시 금속판재에 코팅된 필름의 거동에 대해 분석하여 VCM 소재의 성형 한계에 대한 연구와 [1], 소재 성형 시 필름과 금속 간 접착 거동에 대해 실험적, 해석적으로 규명 [2],[3]하는 등 다양한 방식으로 원인 분석과 개선방안을 위한 연구가 진행되었다.

하지만 소재의 종류, 제품의 형상, 금형의 상태에 따라 조건이 달라지므로 궁극적인 해결 방안은 되지 못하며 여전히 보호 비닐에 의한 찍힘 불량은 빈번하게 나타나고 있다. 외관 불량을 유발하는 다른 원인으로는 칩에 의한 찍힘 혹은 스크래치이다. 칩이란 금속 부스러기를 말하며 이는 전단 성형 시 주로 발생하며 금속과 소재 간 마찰에 의해서도 발생된다. 전단 성형이란 금형을 이용하여 소재를 잘라내거나 구멍을 뚫는 등 절단하는 성형을 말한다. 금형을 이용하여 소재를 전단 하는 과정에서 가공되는 소재와 금형의 마모 상태, 금형의 구조 등에 의하여 칩이 발생하며 이 칩은 보호 비닐의 찌꺼기와 같이 금형 혹은 소재 위에 남아 찍힘과 스크래치 불량을 유발한다 [그림 6]. 이 외에도 피가공 소재와 금형 간 마찰에 의해 긁힘이 발생하거나 여기에서 발생한 이물이 남아 불량을 야기시키기도 한다.

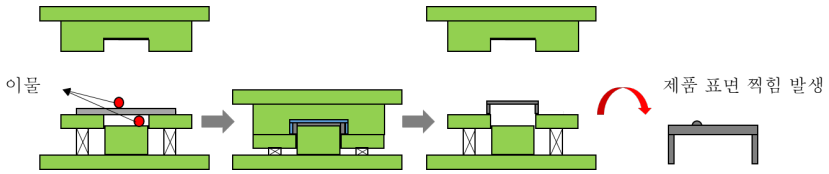


그림 6: 금형 내 이물에 의한 표면 찍힘 발생 과정

프레스 성형은 여러 개의 공정이 연속적으로 이루어진다. 생산성 향상을 위해 공정과 공정 간 소요시간을 줄이기 위해 소재가 연속으로 투입되는데 이물이 금형에 남아있을 경우 뒤이어 투입된 소재에까지 영향을 미치게 되어 연속 불량을 초래한다. 생산 속도가 검사 속도보다 빠를 경우 성형이 완료된 제품을 적재 후 검사를 하게 되면 불량 원인 제거가 늦어져 다량의 불량이 발생되게 된다. 이에 생산 완료와 동시에 검사를 수행하여 연속 불량을 억제해야 할 필요가 있다.

2.2 머신 비전

머신 비전은 고성능 카메라와 디지털화 된 데이터의 처리 정치로 구성된 영상 장치를 말하며 측정된 이미지의 영상 처리를 통해 객체의 유무 혹은 이상 감지, 패턴, 문자 분석 등 다양한 분야에 적용되고 있다. 생산성 향상을 위한 자동화 검사를 위해서는 전문 검사원의 육안 검사를 대신하는 고속, 고해상도의 정밀검사가 가능해야한다.

머신 비전은 인간의 시각을 기계가 대신 하도록 하여 제품 생산 시 결함 분석, 형상, 표면 품질 검사 등에 사용되고 있으며 머신 비전 기술의 발전으로 반도체, 전자 부품, 의료, 자동차 등 다양한 산업 분야에서 활용되고 있다. [4]

머신 비전 시스템은 표면의 결함, 이물질, 문자, 위치 등 다양한 분야에서 활용되며 [그림 7]과 같이 일반적으로 카메라, 렌즈, 조명 등의 광학계와 영상처리 장치, 프로그램 및 시스템 제어 장치로 구성된다. 조명으로부터 발생된 빛이 물체로부터 반사되어 렌즈를 통해 카메라로 전달 후 영상처리장치를 통해 디지털 신호로 변환된 신호가 소프트웨어의 알고리즘을 통해 결과가 출력되는 시스템이다. 대상 물체의 보다 정확한 정보를

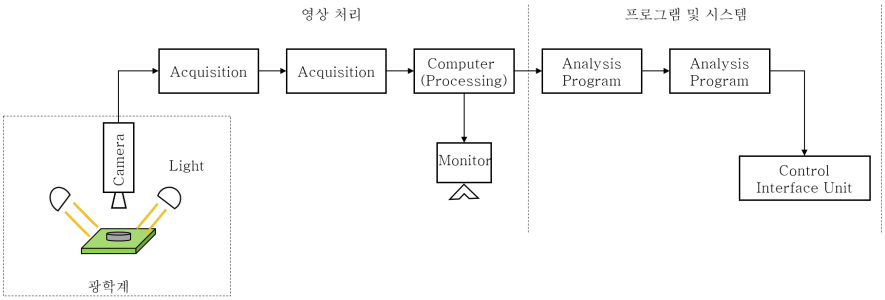


그림 7: 머신 비전 시스템 구성[5]

얻기 위해 카메라의 구성, 수량, 위치 변화 등 카메라의 변화에 따라 오차를 줄이는 방법 [6], 알고리즘의 개선을 통해 대상물의 인식 정확도 향상 [7], 조명의 밝기, 수량, 위치의 변화를 [8] 통해 더욱 정확한 검사가 가능한 방안들에 대한 연구가 진행되어 왔다.

이 중 특히 조명은 대상 물체의 상태를 알아내는 데에 매우 중요한 요소이다. 주변 환경과 피측정물에 적합한 조명을 구성함으로써 데이터의 신뢰도와 정확도 향상, 검사 시간을 단축시킬 수 있다. 불량인지 양품인지, 어떤 문자가 있는지, 혹은 어떤 표면을 가지는지 등 머신 비전 시스템에서의 시작인 이미지 획득 부분에서 조명이 적합하지 않을 경우 원하는 결과를 얻기 어렵게 된다. [9]

2.2.1 조명

머신 비전에서 조명은 검사 대상인 물체의 그림자 및 빛의 반사를 최소화하여 물체가 가진 특성을 영상화 하기 위한 요소로 머신 비전 시스템, 특히 이미지를 통한 실시간 자동화 검사 환경을 구축할 시 매우 중요한 요소이며 조명의 종류, 밝기, 형태 등 다양한 고려사항들이 있다.

조명은 조사하는 방식에 따라서 후면광(Back light unit), 확산광(Dif-

fuse light unit), 명시야(Bright field unit), 암시야(Dark field unit) 등으로 분류된다. [그림 8]과 같이 후면광은 대상 물체의 후면에서 조명을 비춰 검사하는 방식이며 보통 물체의 위치, 방향, 높이 등을 검출하는데 사용된다. 확산광은 반사율이 높으며 반구형의 돔 형태 또는 동축으로 조명이 설치된다. 돔 형태의 조명은 반사율이 높은 물체, 동축으로 사용 시 평면 형태의 물체에 적합하다.

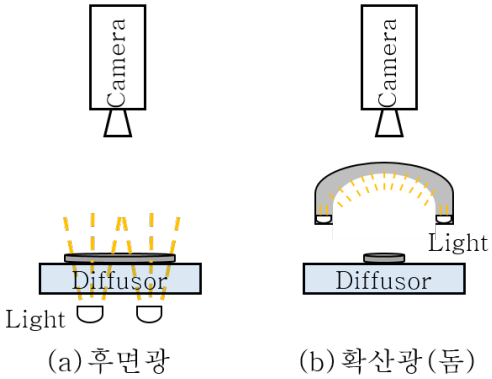


그림 8: 후면광과 확산광의 조명 조사 방식

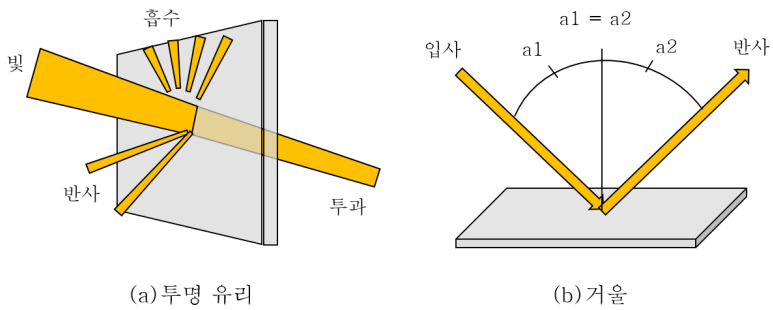


그림 9: 투명유리와 거울면에서의 빛의 반응

[그림 9]는 투명한 유리와 거울면에서의 빛의 상태를 나타낸다. 투명

한 유리의 경우 빛은 투과, 반사, 흡수, 발산으로 나뉘며 전체 조사광과의 총합이 같다. 거울의 경우는 평면에서의 빛의 입사각과 반사각의 크기가 항상 같다. 투명한 유리, 거울면과 같이 조사된 빛이 그대로 반사된 빛을 직접광이라 하며, [그림 10]과 같이 전방위로 산란된 빛을 산란광이라 한다. 조명의 방법에는 이러한 빛의 반사를 통해 반사율과 투과율의 명암 차이를 이용하여 밝은 이미지를 얻는 직접광 조명 방법, 산란율의 명암 차이를 이용하여 어두운 이미지를 얻는 산란광 조명법이 있다. 직접광은 밝은 이미지를 얻을 수 있기 때문에 명시야 조명, 산란광은 암시야 조명이라고도 불린다. [그림 11]은 조명의 위치에 따른 암시야와 명시야의 구분을 나타낸다.

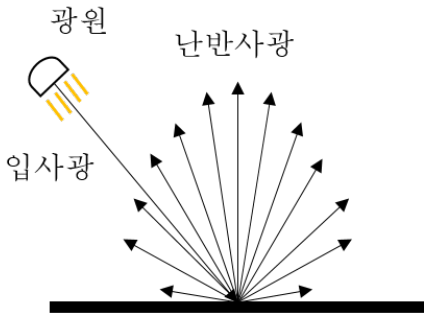
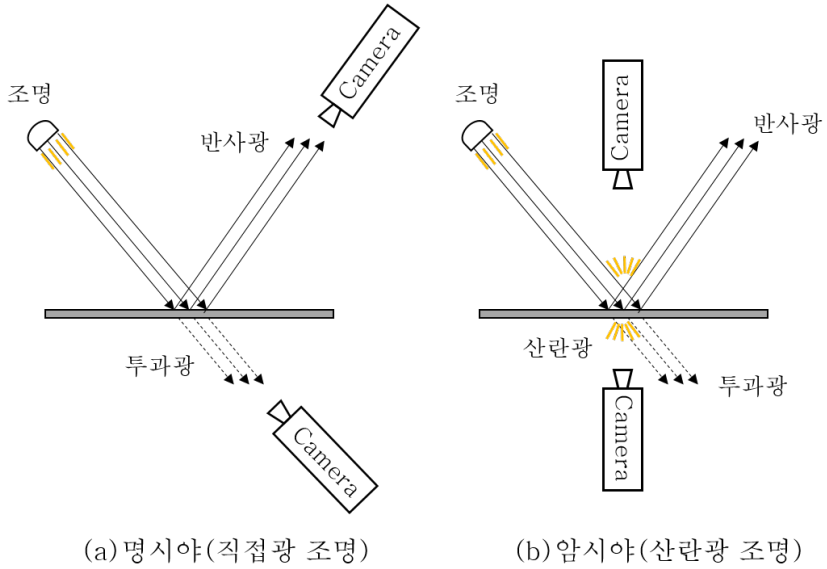


그림 10: 산란광[10]

조명은 광원에 따라 LED, 할로겐, 형광, 제논 등으로 구분하며 일반적으로 깜박임이 없고, 단색광, 장수명, 안전성 등 사용이 용이한 LED 조명이 많이 사용되고 있으며 이 외에도 적외선, 자외선 조명, 레이져 등 물체의 재질, 형상 및 주변 환경에 적합한 조명이 활용되고 있다. 물체의 표면 상태와 조명의 형태에 따라 빛을 조사하는 방식과 얻어지는 이미지가 달라지는데 그 차이들과 이를 어떻게 처리하여 검사에 활용하는지는 추 후 언급하고자 한다.



(a) 명시야(직접광 조명)　　(b) 암시야(산란광 조명)

그림 11: 조명의 위치에 따른 암시야와 명시야

이와 같이 조명의 종류, 반사 방법, 형태 등에 따라 얻게 되는 이미지는 크게 달라질 수 있으며 대상 물체의 특성, 형상, 표면 상태, 반사의 형태와 정도, 주변 환경 등을 고려하여 최적화 된 조명을 선정해야 한다. 특히 본 연구의 대상 물체인 반사 특성이 높은 부품의 경우에는 더욱 주의해야 한다.

2.3 딥러닝과 합성곱 신경망

2.3.1 딥러닝

기계로 인간의 지능, 특히 학습, 추론, 지각의 능력을 인간과 유사하게 구현하고자 하는 것이 인공지능(AI:Artificial Intelligence) 기술이며 인공지능을 구체화하기 위한 방법으로 인공지능의 하위 개념인 머신러닝,

딥러닝 기술이 발전되어 왔다. 인공지능, 머신러닝과 딥러닝은 [그림 12]과 같은 범주로 구분된다.

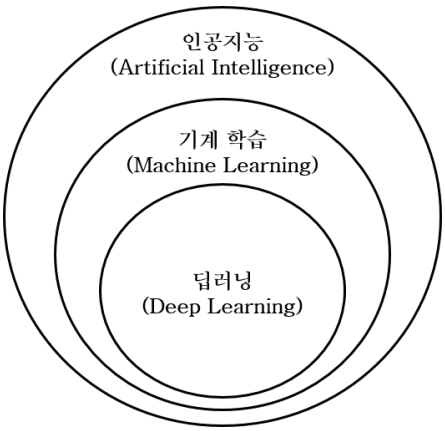


그림 12: 인공지능, 기계학습, 딥러닝의 벤다이어그램

머신러닝은 기계가 일일이 명시하지 않은 기능을 데이터를 분석하고 학습하여 판단, 예측 등의 실행을 할 수 있도록 하는 알고리즘을 개발하는 분야이다. 딥러닝은 머신러닝의 한 분야로서 인공신경망을 기반으로 대량의 데이터의 패턴, 규칙등을 스스로 추출하여 학습하여 뛰어난 성능을 이끌어내는 연구 분야이며 최근 많은 양의 데이터(Big data), 머신러닝 알고리즘과 GPU 성능의 향상으로 빠른 발전을 이루고 있다.

딥러닝은 인공신경망 기반의 모델로 인간의 뇌신경 뉴런의 시냅스 간 결합 강도에 따른 관계를 수학적으로 모델링한 Flank Rosenblatt 의 퍼셉트론 [11]을 기반으로 발전된 기술이다. 인공 신경망은 [그림 13]와 같이 데이터를 받는 입력층, 다수의 노드와 레이어를 통해 특징 추출을 위한 은닉층, 결과를 출력하는 출력층으로 구성되며 딥러닝은 이러한 인공 신경망을 깊게 쌓아 학습하는 모델이다 [12].

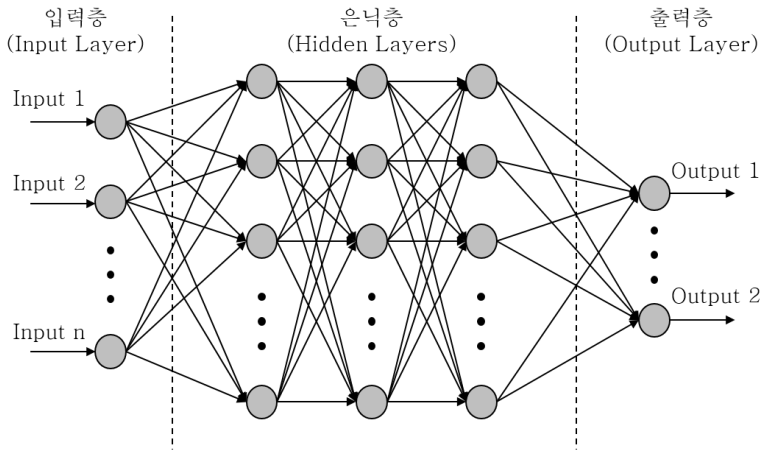


그림 13: 인공신경망

딥러닝 모델은 실제값과 예측값의 오차를 최소화하는 인자를 찾는 알고리즘이다. 학습과정을 설명하면 순전파(Feedforward)를 통해 입력된 데이터를 각 층에서 가중치와 활성화 함수(Activation) 연산을 반복하여 예상 결과값을 출력한다. 순전파 과정에서 데이터의 특성이 n개일 때 다음 식 (2.1)으로 정의되며 w는 가중치(Weight), b는 편향(Bias)를 의미한다.

$$y = \sum (b + w1x1 + w2x2 + ... + wnxn) \tag{2.1}$$

이 후 출력층에서 예측된 출력값과 입력된 데이터의 차이를 손실 함수(Loss function)를 통해 계산한다. 손실 함수란 입력 데이터가 순전파를 거쳐 출력층을 통해 출력된 값과 입력데이터와의 차이를 계산하는 과정이며 손실 함수의 값이 높을수록 성능이 낮다는 것을 의미하므로 신경망은 이 손실 함수가 최소화 되도록 가중치를 갱신하는 것이다. 손실함수로는 평균제곱오차(Mean Squared Error), 이진 교차 엔트로피(Binary

Cross-Entropy), 범주형 교차 엔트로피(Categorical Cross Entropy)가 주로 쓰인다. 교차 엔트로피 손실은 분류 문제에 많이 사용되는 손실 함수이며 식 (2.2)로 표현된다. 여기서 i는 데이터셋의 i번째 데이터, y_i는 i번째 데이터의 실제 값, $\hat{y}$은 i번째의 데이터에 대한 예측값이며 예측 확률이 실제 값에서 멀리 떨어질수록 증가하는 형태를 보인다. 교차 엔트로피는 분류해야 할 클래스가 2보다 클 때 사용하며 이진 교차 엔트로피 손실함수는 클래스가 0과 1 인 이진 분류일 때 사용하며 식 (2.3)과 같이 표현된다.

$$CrossEntropyLoss = -(y_i log(\hat{y_i}) + (1 - y_i)(log(1 - \hat{y_i})) \tag{2.2}$$

$$BinaryCrossEntropyLoss = -\frac{1}{N}\sum_{i=1}^{N}(y_i * ln\hat{y_i} + (1 - y_i) * ln(1 - \hat{y_i})) \tag{2.3}$$

본 연구는 양품과 불량을 분류하는 이진 분류 문제로 이진 교차 엔트로피를 손실함수로써 사용하였다. 가중치의 갱신은 경사하강법(Gradient descent)과 역전파(Backpropagation)를 통해 이루어진다. 역전파란 순전파와 반대 방향인 출력층에서 입력층으로 각 가중치를 얼마나 갱신해야 하는지 구하는 알고리즘으로 딥러닝에 있어 핵심적인 역할을 한다 [13]. 역전파는 경사하강법 [14]을 통해 오차가 최소화되는 방향을 계산하여 진행하는데 이를 최적화라 하며 이와 같은 일련의 과정을 반복하여 최종 출력값과 입력데이터의 오차를 최소화하는 방향으로 모델을 학습시킨다.

딥러닝의 대표적인 모델인 다층 퍼셉트론의 경우 데이터를 1차원 형태로 사용하기 때문에 이미지와 같은 2차원 데이터가 입력 시 1차원의 열벡터로 변환한다. 이 경우 이미지 데이터의 정보와 특징, 특히 공간 정보가 무시되며 전체 픽셀에 대한 가중치 계산으로 인해 많은 학습시간이 필요하고 과적합(Overfitting)의 우려가 있다. 이를 해결하기 위한 방안으로 데이터의 공간적 특징을 포함하며 효율적으로 데이터의 특성을 추출할 수 있도록 하는 합성곱 신경망(CNN)이 고안되었다.

2.3.2 합성곱 신경망

합성곱 신경망(CNN:Convolution Neural Network)은 영상과 이미지 등 컴퓨터 비전 분야의 대표적인 인공신경망 모델 중 하나로 1989년 LeCun이 필기체의 우편번호를 인식하여 분석하는 연구로부터 처음 시작되었다[15]. 이 후 2003년 Behnke에 의해 일반화 [16], Simard에 의해 단순화[17] 되면서 합성곱 신경망이 발전해오게 되었다. 합성곱 신경망은 데이터를 1차원 평면화 시켜 공간 정보를 잃게 되는 DNN과 달리 이미지 데이터의 공간적, 지역적 정보를 유지한 채 특성 추출이 가능하여 컴퓨터 비전 분야에서 좋은 성능을 보인다.

합성곱 신경망의 대표적인 구조는 [그림 14]와 같으며 데이터 입력 후 크게 영상의 특성을 추출하는 부분과 분류하는 부분, 두가지로 구분될 수 있다. 특징을 추출하는 부분에서는 합성곱층(Convolutional layer), 풀링층(Pooling layer)이 있으며 분류하는 부분에는 전연결층(Fully connected layer)과 출력층(Output)이 있다. 입력층을 통해 이미지 데이터가 입력되면 합성곱층에서 각각의 가중치를 가진 필터로 합성곱 연산을 수행하여 이미지의 특징을 추출한 후 특징맵(Feature map)을 생성하며 그 과정은 [그림 15]와 같다. 합성곱층은 Filter 크기, 이동 간격(Stride), Padding 크

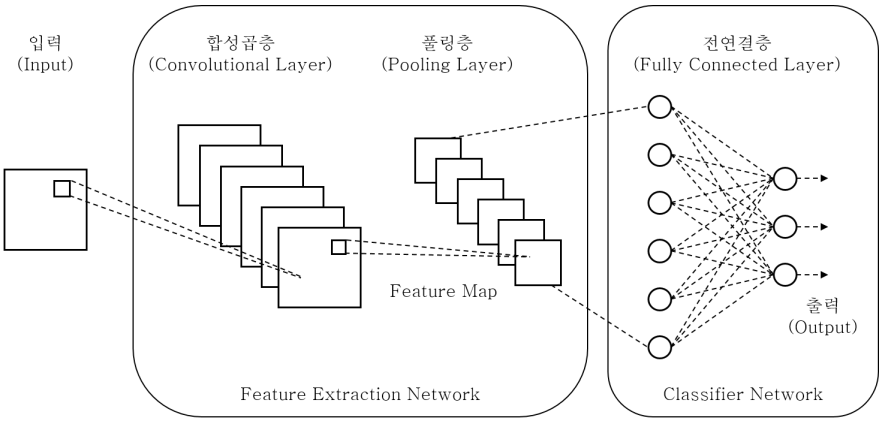


그림 14: 합성곱신경망의 구조

기에 따라 출력 데이터의 Shape이 달라진다. 합성곱층 다음에 위치하는 풀링층은 서브샘플링(Sub-sampling)이라 표현되기도 하며 선택적인 층이다. 입력된 이미지가 합성곱을 거치면서 크기를 유지한 채 다음층으로 가게 되면 연산량이 기하급수적으로 많아지게 되므로 연산 속도가 증가되고 과적합(Overfitting)이 발생할 우려가 있다. 이를 방지하기 위하여 풀링층에서 파라미터(Parameter)를 줄여 처리 속도를 향상시키고 특징을 강조시키는 역할을 한다. [그림 16]과 같이 풀링은 최대값(Max pooling), 평균값(Average pooling), 최소값(Min pooling)이 있으며 주로 각각의 패치내 가장 큰 값을 선택하는 최대값을 사용한다.

정리하면 특징을 추출하는 부분에는 합성곱층과 풀링층이 있으며 합성곱층은 이미지의 특성 추출, 풀링층은 이미지의 크기를 줄이는 역할을 하며 [18] 두 층을 반복적으로 사용하여 특징을 찾아 전연결층으로 보낸다. 전연결층은 앞서 설명한 인공신경망과 같이 전 노드가 완전 연결(Fully connected) 된 층으로서 합성곱 층으로부터 추출된 특징맵을 입력 받아 3차원의 이미지 벡터를 1차원의 열벡터로 변환하여 소프트맥스(Softmax),

시그모이드(Sigmoid)와 같은 분류기를 통해 분류하는 역할을 한다.

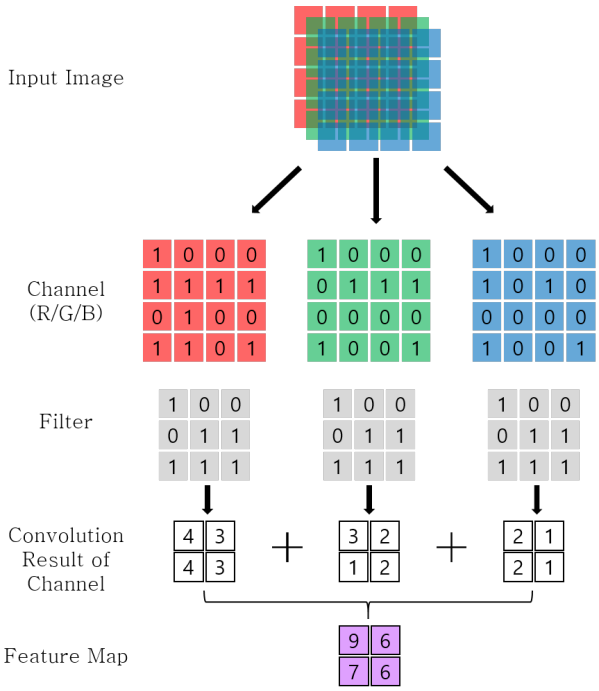


그림 15: 합성곱신경망의 특징맵 추출 과정[19]

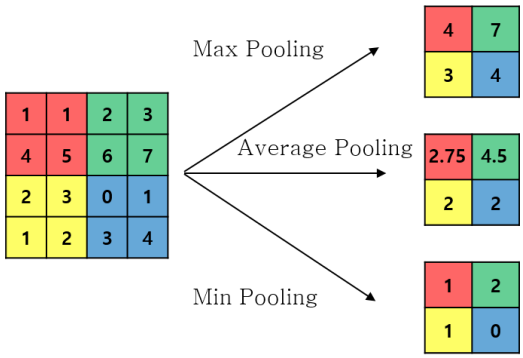


그림 16: 풀링의 종류 별 예시

2.4 합성곱 신경망 아키텍처

합성곱 신경망의 연구는 2012년 딥러닝 기반 이미지 분류에서 우승한 AlexNet을 시작으로 더욱 좋은 결과를 내기 위하여 다양한 연구들이 진행되어 왔으며 높은 성능을 내는 아키텍처들이 개발되었다. 본 절에서는 본 연구보고서에 사용되는 VGGNet, ResNet, Inception, Xception에 대해 간략히 설명한다.

2.4.1 VGGNet

VGGNet[20]은 옥스포드의 Karen Simonyan과 Andrew Zisserman이 개발하여 2014년 ILSVRC 대회에서 2 위를 차지한 모델이다. 당시 1위인 GoogLeNet [21] 보다 성능은 낮았지만 구조가 간단하고 단일 네트워크에서 좋은 성능을 보여 최근에는 이미지의 특징추출에 이용되는 등 기본 네크워크 모델로 활발히 활용되고 있다. VGGNet의 기본 구조는 [그림 17]과 같다.

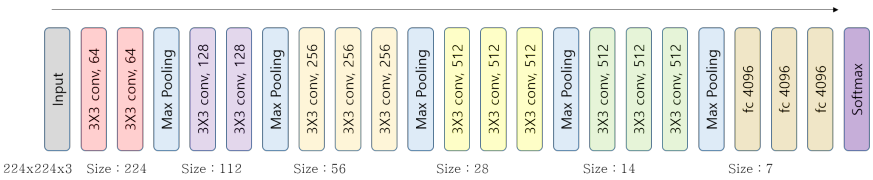


그림 17: VGGNet 구조

VGGNet의 가장 큰 특징 중 하나는 이전 다른 네트워크들이 7x7, 11x11등 비교적 큰 필터들을 사용한 것과 달리 3x3 필터를 여러 개 사용하는 것이다. Stride를 1로 했을 경우 풀링 없이 3x3을 두 번 사용하는

것과 5x5를 한번 사용하는 것이 같고, 7x7의 경우 3x3 필터를 세번 사용하는 효과와 같다. 세번의 비선형 함수가 적용됨에 따라 모델의 식별성 증가의 효과가 있으며 C개의 채널을 가진 7x7 필터를 한 번 사용 시 파라미터 수는 다음과 같이 $7^2\ C^2$=$49C^2$, C개의 채널을 가진 3x3 필터를 세번 사용 시 파라미터수는 $3(3^2\ C^2)$=$27C^2$로 파라미터의 수가 81% 감소하는 효과가 있다.

VGGNet 논문의 정확한 이름은 Very Deep Convolutional Networks For Large-scale Image Recognition으로 층을 깊게 쌓아 높은 성능을 낼 수 있음을 보여주었지만 단순한 구조와 좋은 성능에도 불구하고 마지막단의 전연결층으로 인해 파라미터의 개수가 너무 많아지며 기울기소실과 폭주, 과적합등의 문제가 발생할 수 있다는 단점이 있다.

2.4.2 ResNet

VGGNet 논문을 통해 층이 깊어질수록 성능이 더 좋다는 것을 확인할 수 있었지만 층이 너무 깊어지면 기울기 소실, 폭주와 같은 성능 저하 문제(Degradation problem)의 우려가 있다. 기울기 관련 문제를 해결하는 방안으로 층 중간에 배치정규화(Batch normalization) 적용 [22], 네트워크의 가중치 값의 적절한 초기화 등의 테크닉 등이 제안되었다 [23]. 성능 저하 문제란 층이 깊으면 무조건 정확도가 높아지는 것이 아닌 너무 깊을 경우 정확도가 오히려 감소하는 문제로 에러가 높아지는 현상이다. 이를 해결하기 위한 방안으로 2015년 Kaiming이 Deep Residual Learning Framework, 줄여서 ResNet을 제안하였다 [24].

기존 인공신경망의 학습 목적은 인풋(x)를 타겟(y)로 맵핑하는 함수 H(x)를 찾는 것이므로 $H(x) - y$를 최소화하는 방향으로 학습을 진행하는 것이었다면 ResNet은 $H(x) - x$를 새로운 함수 F(x)로 정의하여 F(x)=0이

될 수 있도록 하는, 즉 $H(x)-x$를 최소화하는 함수F(x)를 찾는 것을 목적으로 한다. ResNet에서 잔차(Residual)는 예측값과 실제값의 차이인 잔차를 뜻하며 $F(x)=H(x)-x$ 라고 표현되고 이 잔차를 학습하는 것을 Residual Learning 이라 한다.

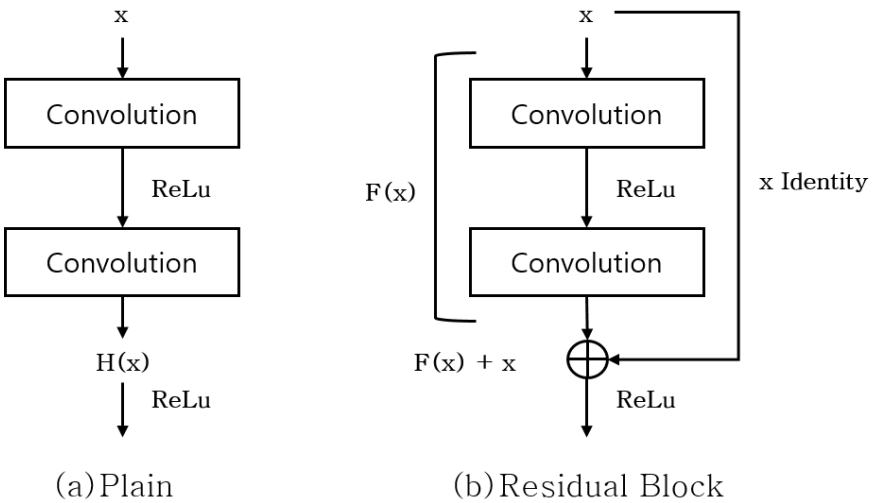


그림 18: Residual Block[24]

[그림 18]의 좌측과 같이 네트워크의 아웃풋이 x가 되도록 맵핑하는 것이 아닌 우측의 잔여 블럭(Residual block)과 같이 네트워크의 아웃풋은 0이 되도록 학습시키고 마지막에 x를 더해서 H(x)가 x가 되도록 학습한다. x가 F(x)를 통과한 후, 자기 자신인 x를 층을 스킵해서 더해주는 것을 Skip Connection이라 부른다. 잔여블럭을 통과하고 난 후의 최종 아웃풋은 $F(x)+x$가 된다. 이 경우 미분을 해도 x 자체의 미분값 1을 가질 수 있기 때문에 각 층마다 그래디언트를 최소 1은 갖도록 하여 기울기 소실의 문제를 해결할 수 있게 된다. 각각의 잔여 블럭은 2개의 3x3 합성곱층을 사용하며 네트워크의 층이 깊어질수록 시간과 파라미터수가 증가하는 문

제를 해결하기 위해 일부 층의 합성곱층을 변경하여 차원을 감소시키는 병목 구조를 적용하였다. ResNet은 기본 구조인 VGGNet에 잔여블럭과 병목 구조를 구성하여 깊은 네트워크에도 기울기 소실 문제의 해결과 빠른 속도로 좋은 성능을 보였다.

2.4.3 Inception(GoogLeNet)

2014년 Szegedgy가 Inception이라는 개념을 제안하였으며 [21] 이를 바탕으로 구현된 GoogLeNet은 2014년 IRSVRC에서 1등을 차지하였다. 앞서 VGGNet의 경우와 같이 네트워크가 깊어질수록 성능이 좋아질 수 있다는 것은 확인하였으나 마찬가지로 많은 층을 쌓을수록 그만큼 파라미터의 수가 많아지고 이에 따른 연산량 증가, 기울기소실, 과적합의 문제가 발생할 수 있다는 우려가 있다. 이를 해결하기 위해 GoogLeNet은 Inception Module을 통해 연산량을 최소화 하면서 모델의 크기를 키우는 방안을 제시하였다. Inception Module은 [그림 19]와 같은 구조로 되어있다.

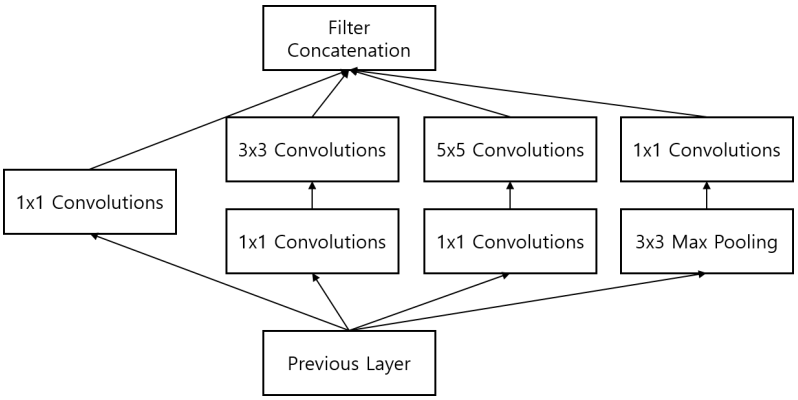


그림 19: Inception Module[21]

인풋에 대해 1x1, 3x3, 5x5의 합성곱층을 병렬로 연결하며 파라미터를 줄이기 위한 방법으로 각 층에 1x1 합성곱층인 병목층을 추가하여 차원 축소와 전연결층의 사용을 최소화 하였다. 또한 기울기소실 문제 해결을 위해 중간층에서 분류 손실(Classification loss)계산을 통한 역전파를 수행하여 가중치 갱신을 할 수 있도록 하였다. 2014년 2번째로 좋은 성능을 보였던 VGGNet은 GoogLeNet 대비 파라미터수와 연산량이 많지만 단순한 구조가 장점으로 이 후 많이 사용되었다면 GoogLeNet은 구조가 복잡하여 네트워크의 변경이 어려운 단점이 있었다. 이에 성능 향상과 연산량 감소를 위해 2015년 Inception v2와 Inception v3가 제안되었다 [25]. Inception v2는 연산량을 줄이면서 층을 깊게 쌓기 위하여 기존의 5x5 필터를 3x3 필터 2개로 대체하였으며 이 3x3 필터를 비대칭적인 1x3 혹은 3x1 필터 2개로 한번 더 분해하여 파라미터 개수를 줄였다. 또한 특성맵이 가진 정보의 손실을 최소화하면서 효율적으로 이미지 사이즈를 줄이기 위하여 [그림 20]과 같이 Inception Module 내에서 풀링과 필터에 Stride 2를 사용해서 크기와 계산량을 함께 줄였다.

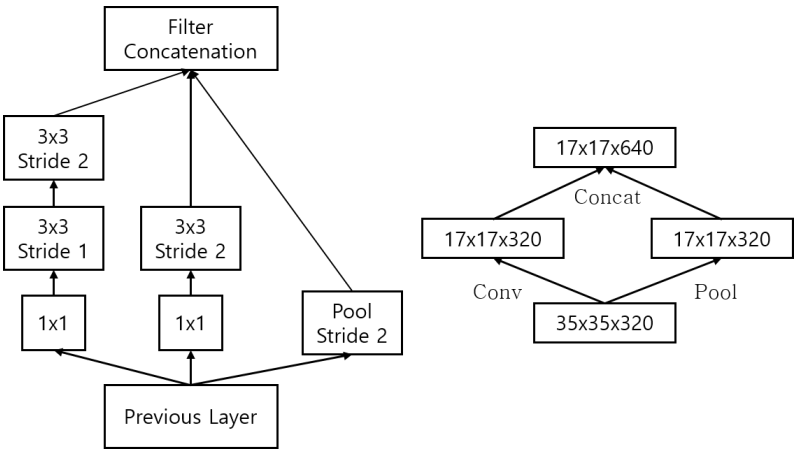


그림 20: Inception V2[25]

Inception v2는 총 42개의 층으로 구성되며 기존 GoogLeNet 대비 파라미터가 2.5배 늘어났지만 VGGNet 대비 효율적인 계산량을 보여주었다. Inception V3는 모델 구조는 Inception v2와 동일하며 여기에 Optimizer를 SGD에서 RMSprop으로, 7x7 필터를 3x3 필터 3개로 변환, 마지막 전연결층에서 Batch Normalization을 적용하는 BN-auxiliary가 추가되어 성능을 개선하였다.

2.4.4 Xception

Xception [26]은 구글이 2017년 발표한 모델로 Inception v3보다 좋은 결과를 냈다. Inception 모델들은 Inception Module을 쌓아 올려 적은 파라미터로 좋은 성능을 낼 수 있도록 고안된 Module로, 앞서 설명한 바와 같이 1x1, 3x3, 5x5로 구성되어 있으며 1x1 합성곱층은 Cross Channel Correlation을, 3x3과 5x5는 Spatial Correlation을 담당한다. 이로써 Channel과 Spatial Correlation을 분리하므로 적은 파라미터로 높은 성능을 낼 수 있게 되었다. Inception Module을 Channel Correlation과 Spatial Correlation을 분리하면 성능이 더 좋다는 가정하에 Inception Module을 [그림 21]과 같이 단순화된 구조로 볼 수 있게 되며 이를 더 단순화하여 1x1 합성곱층을 두고 이 층으로부터 출력된 아웃풋에 대해 겹치지 않게 사용되는 Spatial 합성곱층을 사용한다고 했을 때 [그림 22]와 같은 구조를 생각할 수 있게 된다.

Inception 모델에서 Channel Correlation과 Spatial Correlation를 잘 분리하여 좋은 성능을 내는 것을 확인하였다면 Xception 모델은 이를 완벽히 분리하여 독립적으로 계산하고 맵핑하기 위해 고안된 모델이다. 이를 Extreme Version of Inception Module로 Xception이라 부르며 Inception v3 대비 더 좋은 성능을 낼 수 있게 되었다.

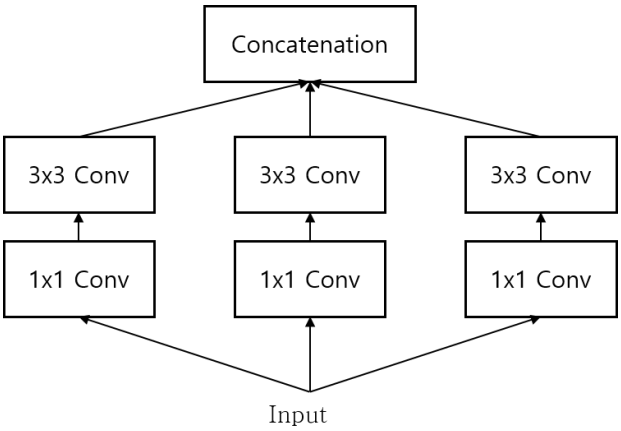


그림 21: 단순화 한 구조의 Inception Module[26]

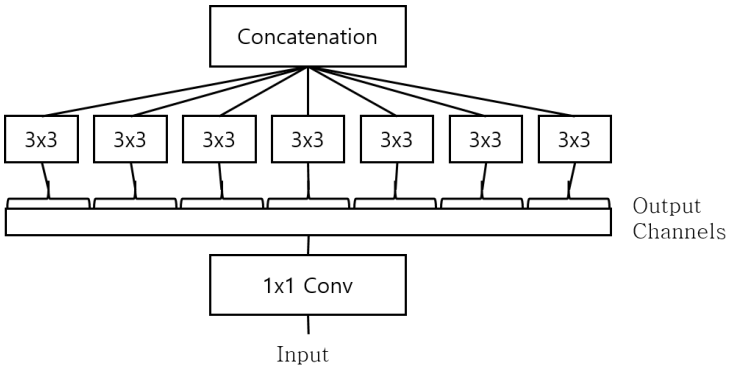


그림 22: Xception[26]

2.4.5 EfficientNet

EfficientNet [27]은 2019년 Mingxing Tan이 제안한 모델로 제한된 리소스(Resource)범위에서 기존 합성곱 네트워크보다 8.4배 적으면서 6.1배 빠르고 더 높은 정확도를 보였다. 그동안 모델의 크기를 키워 성능을 높이는 Scaling up에 대한 연구가 많이 이루어 졌으며 [그림 23]와 같이 그 방법으로 세가지가 있다. 첫번째는 채널의 넓이를 늘리는 것, 두번째는 채

널의 깊이를 깊게 만드는 것, 그리고 마지막으로 인풋 이미지의 해상도를 높이는 것이다.

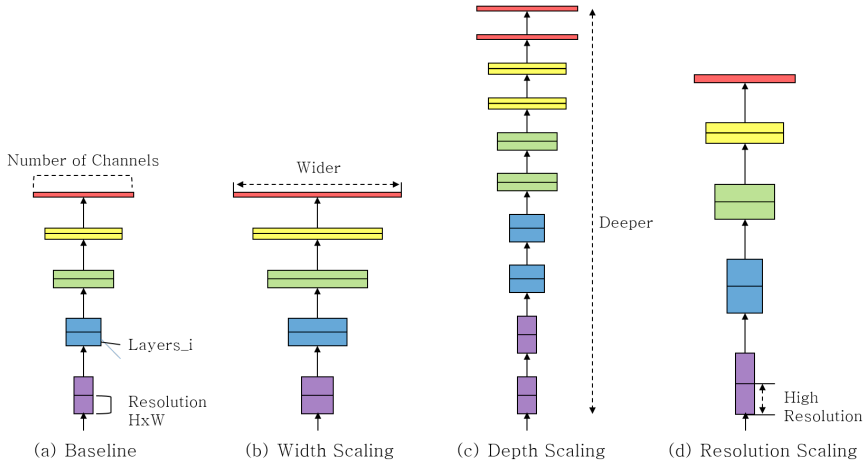


그림 23: Scaling Up 세가지 방법[27]

네트워크의 깊이를 깊게 만들면 좀 더 복잡한 특징을 추출할 수 있으나 기울기 소실 발생의 우려가 있으며 채널의 폭을 증가 시킬수록 세밀한 특징을 찾아내기 용이하나 복잡한 특징은 파악하기가 힘들다. 마지막으로 고해상도의 이미지를 사용할 경우 세밀한 특징을 잡아내어 정확도를 높일 수 있지만 해상도가 증가할수록 성능 증가폭은 감소되는 단점이 있다.

EfficientNet은 AutoML을 통해 이 세가지 Scaling 방법에 대한 최적의 조합을 찾았으며 효율적으로 조절 가능한 Compounding Scaling 방법을 제안하였다. Compounding Scaling의 기본 개념은 깊이, 넓이, 해상도 등 세가지의 변수는 밀접하게 연관되어 있으므로 단일 차원의 Scaling이 아닌 복합적인 차원의 균형과 조정이 될 수 있도록 하는 것이다. Compound

Scaling은 아래 식 (2.4)와 같이 구성된다. 여기서 ø 는 자원이 추가될 경우 Model Scaling을 비례하여 증가시킬 계수이며 연산량인 FLOPs(FLoating point Operations Per second)는 깊이에 비례하며 나머지 두 변수에 대해서는 제곱에 비례한다. 논문에서는 FLOPs를 ø에 따라 2배씩 증가하도록 설정하기위해 $(\alpha \cdot \beta^2 \cdot \gamma^2) = 2$로 제한하였다. 깊이와 해상도에 제곱항이 있는 이유는 깊이가 2배 증가하면 FLOPs는 2배가 되지만 다른 두개의 변수는 그렇지 않기 때문이다.

$$
\begin{aligned}
depth : d &= \alpha^{\phi} \\
width : w &= \beta^{\phi} \\
resolution : r &= \gamma^{\phi} \\
s.t. \alpha \cdot \beta^2 \cdot \gamma^2 & \\
\alpha \geq 1, \beta \geq 1, \gamma \geq 1 & \qquad (2.4)
\end{aligned}
$$

EfficientNet은 한정된 자원으로 Compound Scailing을 통해 깊이, 넓이, 해상도를 효율적으로 조정하여 크기와 연산량을 줄이면서 빠른 속도와 높은 정확도를 보였다.

2.5 전이학습

학습 성능을 높이는 방법에는 여러 종류가 있으며 그 중 가장 중요한 것 중 하나는 충분한 데이터의 양 일 것이다. 보유한 데이터의 수량이 충분히 많지 않은 경우 학습을 진행하더라도 좋은 결과를 얻을 수 없다. 산

업현장에서 양품과 불량을 구분하기 위한 데이터를 모은다고 가정할 때 양품과 불량의 데이터 중 양품의 수가 압도적으로 차이가 많다. 불량에 대한 데이터가 충분히 모인다 라는 것은 이미 대량 생산으로서 이점이 없기 때문에 양품과 불량을 구분할 필요가 없다는 것을 의미한다. 일반적으로 많이 사용하는 개나 고양이의 분류 문제와 같이 쉽게 얻을 수 있는 이미지와 달리 산업현장에서 데이터의 수집, 특히 학습하기 충분한 불량데이터를 얻기에는 많은 비용과 시간이 필요하다. 이에 대한 방안으로 전이 학습 [28]이 사용된다.

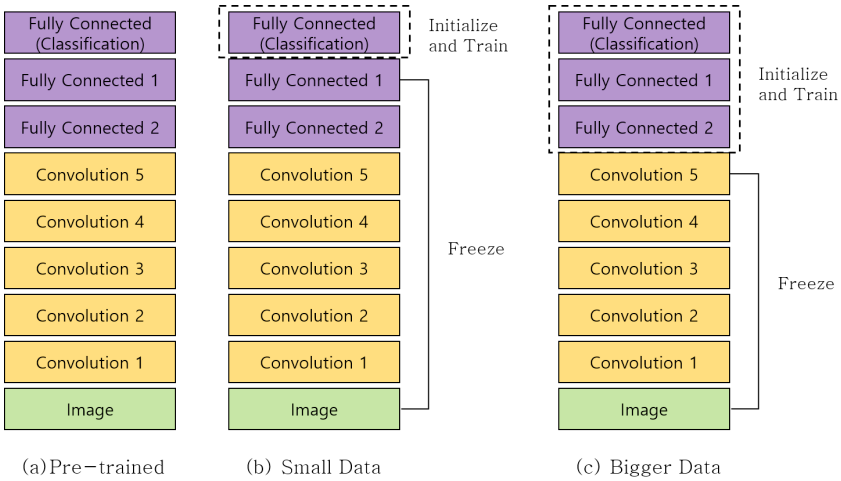


그림 24: 데이터 규모에 따른 전이학습 방법 및 구조

전이학습이란 사전 학습된 모델의 가중치를 바탕으로 새로운 모델을 만들어 학습을 빠르게 하는 방식이다. 보유한 데이터의 양과 특성에 따라 사전 학습된 모델의 가중치를 어느 단계에서 적용할지 결정한다. 데이터의 수량이 많은 경우 사전 학습된 전체 모델에 대해 매개변수의 세부 조정을 하는 방법과 데이터 수량이 적을 경우 특징 추출은 사전 학습된

모델의 매개변수를 사용하고 최종 분류기만 학습시키는 방법 등 데이터에 따라 성능의 향상과 학습시간 단축을 위한 다양한 방법들이 적용 될 수 있다. 전이 학습을 사용할 경우 특징 추출을 위한 별도의 학습이 필요하지 않게 되므로 기존 모든 네트워크를 학습하는 딥러닝 기반의 합성곱 신경망 대비 연산량 감소에 따른 학습 시간 단축, 적은 데이터로도 높은 성능을 얻을 수 있다. [그림 24]는 데이터의 규모에 따른 전이 학습의 방법과 구조를 나타낸다.

2.6 포아송 이미지편집

2003년 Patrick Pérez가 제안한 포아송 이미지 편집 방법(Poisson image editing) [29]은 목표 이미지의 관심 영역으로 소스이미지를 합성 경계가 없는 새로운 이미지를 만드는 방법이다. 포아송 이미지 편집은 이차 편미분 방정식인 포아송 방정식을 사용하는 기울기(Gradient) 기반의 영상처리 방법으로 목표 이미지와 소스 이미지의 기울기가 최소화되고 경계인 $\partial\Omega$ 에서 목표 이미지와 소스이미지의 픽셀값을 동일하게 만드는 방법이다. [그림 25]에서 v는 소스이미지의 기울기, g 는 소스이미지에서 선택된 영역, f* 는 목표 이미지의 색상, f는 합성 이미지를 의미하며, Ω 는 g 의 배경이 되는 목표 이미지를, $\partial\Omega$ 는 목표 이미지와 합성 이미지의 경계를 나타낸다.

포아송 이미지 편집은 목표 이미지의 한 영역 Ω에서 소스이미지를 보간(Interpolation)하는 함수 f(x,y)를 찾는 것이라고 말할 수 있으며 f는 보간 영역의 경계인 $\partial\Omega$에서는 목표이미지와 같은 색상 f*를 가져야 한다. Ω를 경계조건으로 $v = \nabla f$ 를 최적화하면 식 (2.5)와 같고 $f' = f - v$라고 정의하면 $v = \nabla f$는 기울기 이므로 식 (2.6)과 같으며 라플라스 방정식으

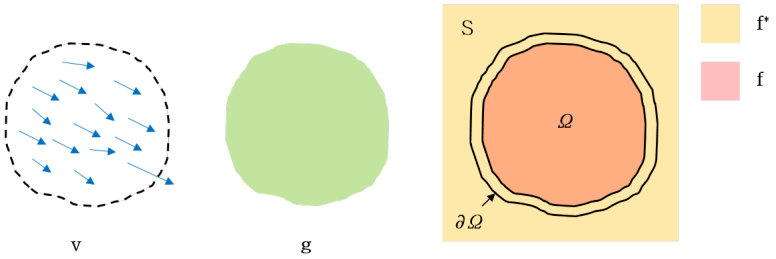


그림 25: 포아송 이미지 편집 기법

로 표현하면 식 (2.7)과 같다. 식 (2.7)과 같이 합성이미지의 경계값과 목표 이미지의 경계값이 같다는 조건과 소스이미지의 한 화소를 4배한 후 주변 4개의 화소값의 합의 차인 라플라시안(Laplacian)을 이용하여 2차 미분값이 일정한 값이 되도록 하는 방정식을 포아송 방정식이라 한다.

$$\min_{f} \iint_{\Omega} |\nabla f - \mathrm{v}|^2 \, with f|_{\partial\Omega} = f^*|_{\partial\Omega} \tag{2.5}$$

$$\min_{f} \iint_{p \in \Omega_0} \left|\nabla f'\right|^2 dp \; with \; f'\left|_{\partial\Omega_0} = (f_t - f_s)\right|_{\partial\Omega_0} \tag{2.6}$$

$$\triangle f' = 0 \; with \; f'\left|_{\partial\Omega_0} = (f_t - f_s)\right|_{\partial\Omega_0} \tag{2.7}$$

포아송 이미지 편집은 합성 경계가 없는 편집에 우수한 성능을 보이지만 합성하려는 배경이미지의 색상이 복잡하거나 목표와 소스의 색상차이가 크면 부자연스러운 편집이 발생할 수 있는 단점이 있다.

제 3 장

문제 정의 및 해결 방안

본 장에서는 본 연구의 목표와 해결해야 할 문제를 명확하게 정의하고 연구 범위 및 문제의 해결 방안에 대해 설명한다. 1장에서 언급한 바와 같이 본 연구의 목표는 고광택 제품의 비정형 미세 결함의 최적 검출 방법을 제시함에 있어 빛 반사가 심한 제품에 대해 비전 센서를 통한 이미지 취득 방법, 데이터 부족 문제 해결, 효율적인 특성 추출 방법 및 결함 검출 정확도 향상 방법을 제시하는 것이다. 1절에서는 본 연구에서 해결해야 할 문제에 대해 설명하고 2절에서는 해결 방안에 대해 설명한다.

3.1 문제 설명

산업이 발전함에 따라 제조, 이송, 검사 등 많은 분야에서 자동화가 이루어져 왔으며 이에 따른 생산성 향상, 인건비 절약 등의 효과를 얻을 수 있게 되었다. 최근에는 딥러닝 기술의 발전에 따라 자동화 가능 범위가 더욱 넓어지고 있다. 하지만 모든 공정이나 제품이 자동화가 가능한 것은 아니며 여전히 수작업이 행해지는 현장도 존재한다. 본 연구는 여러 가지 제약 조건으로 인하여 자동화가 이루어지지 못하는 여러 분야 중 제품 특성에 따라 자동화 검사가 이루어지지 못하는 제품에 대해 검사가 가능할 수 있도록 하는 방안에 대해 제시하고자 한다.

본 연구의 목표인 딥러닝을 통한 고광택 제품의 미세 결함 탐지를 위해 해결해야 할 문제로 첫번째는 데이터의 수집이다. 딥러닝의 적용을

위해서도 필요하지만 우선 검사 자동화를 위해서는 불량이 인식 되어야 한다. 자동화 검사에 많이 사용되는 비전 센서를 이용한 검사는 빛의 반사를 이용하여 물체를 인식한다. 제품 표면의 광택도가 낮거나 불량의 형태와 크기가 일정하고 크다면 물체 즉 제품의 합불 인식이 용이하다. 그러나 고광택 표면에 결함마저 잘 보이지 않는다면 비전 센서로 인식할 수 없기 때문에 검사가 어려울 뿐만 아니라 데이터 기반 딥러닝을 적용하고자 할 때에도 의미있는 데이터를 얻을 수 없다. 따라서 고광택 제품 표면의 비정형 미세 결함을 인식할 수 있는 방법이 필요하다. 두번째로는 데이터의 양이다. 딥러닝 적용을 통해 원하는 성능을 얻기 위해서는 충분한 데이터 수가 필요하다. 하지만 산업현장에서 양품 데이터는 쉽게 얻을 수 있으나 불량의 경우 발생 빈도가 낮고 그 수가 많지 않아 학습에 필요한 데이터의 수집이 어렵다. 세번째로는 딥러닝 모델의 검출 성능이다. 본 연구에 적용하는 제품의 불량률은 약 5%이다. 딥러닝 모델의 양품과 불량 분류 성능이 낮은 경우 자동화 검사로서의 역할을 하지 못하게 되고 역효과가 발생할 수 있다. 따라서 본 연구에서 제안하는 방법이 실효성을 가지기 위해서는 다음에서 설명 될 평가지표에서 의미 있는 수치를 나타낼 필요가 있다.

본 연구는 양품과 불량을 인식하는 이진 분류(Binary classification) 문제이다. 이진 분류 문제에서 혼동 행렬(Confusion matrix)은 [표 1] [30]과 같고 정확도(Accuracy), 재현율(Recall), 정밀도(Precision)는 식 (3.1), 식 (3.2), 식 (3.3)으로 표현된다. 앞서 설명한 바와 같이 현장에서 사용할 수 있는 수준이 될 수 있도록 본 연구에서는 상기 지표에서 97% 이상의 결과를 얻는 것을 목표로 하였다.

표 1: 혼동 행렬(Confusion matrix)[30]

구분		예측	
		Positive	Negative
실제	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

$$Accuracy = \frac{TruePositivie + TrueNegative}{TruePositive + FalsePositive + FalseNegative + TrueNegative} \tag{3.1}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{3.2}$$

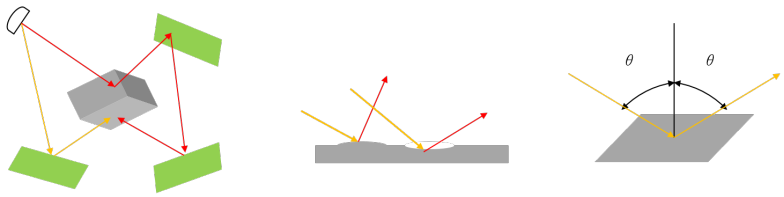
$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{3.3}$$

3.2 해결 방안

3.1 절에서 정의한 문제를 정리하면 빛 반사에 의한 고광택 표면의 미세 결함 인식 난해, 부족한 데이터 수, 현장 사용이 가능한 검출 성능 확보이다. 본 절에서는 해당 문제에 대한 해결 방안을 다음과 같이 제시한다.

첫번째는 비전 센서의 조명 최적화이다. 앞서 설명한 바와 같이 머

신 비전은 조명으로부터 발생된 빛이 물체로부터 반사되어 렌즈를 통해 카메라로 전달 후 영상처리장치를 통해 디지털 신호로 변환된 신호가 소프트웨어의 알고리즘을 통해 결과가 출력되는 시스템이다. 물체는 빛을 받으면 반사를 하며 표면의 상태, 반사의 종류에 따라 복합적인 형태를 보인다. [31]



(a) 주변광(Ambient Reflection) (b) 난반사(Diffusion Reflection) (c) 정반사(Specular Reflection)

그림 26: 반사의 종류

빛 반사는 [그림 26]과 같이 크게 주변광, 난반사, 정반사의 형태를 가지게 되는데 통상적으로 비전 센서를 사용한 이미지를 취득하게 되면 전체 반사율에 대한 정보를 얻게 되므로 표면 상태, 조명 형태에 따라 원하는 정보를 얻지 못하게 되는 경우가 발생한다. 이에 본 연구에서는 조명의 배열과 반사의 형태 중 원하는 정보만 추출하여 표면 상태만 따로 얻을 수 있는 방법을 적용한다.

두번째는 데이터 증량이다. 앞의 머신 비전으로부터 이미지 취득이 가능하지만 여전히 데이터의 수량, 특히 결함 이미지는 부족하다. 적은 양의 데이터를 학습에 충분한 양으로 늘리기 위하여 결함 이미지로부터 결함 부분만 추출하여 양품 이미지에 무작위로 합성하여 결함 이미지를 늘린다. 이 후 회전, 반전 등의 데이터 증량을 통해 학습 가능한 데이터 확보가 필요하다.

마지막으로 검출 성능 향상을 위하여 사전 학습된 모델의 매개변수를 사용하는 전이 학습을 적용하고 추가적으로 특징 추출이 더욱 용이하도록 이미지를 분할하여 학습시키는 방법을 고려한다.

제 4 장

실험 및 결과

4.1 이미지 취득

앞서 설명한 바와 비전 센서는 대상 물체의 반사된 빛을 이용하므로 이미지 취득 시 조명 조건에 따라 결과가 크게 달라진다. 물체가 빛을 받으면 반사가 되고 이 때의 반사 구성 요소, 방향과 전체 반사광은 [그림 27]과 같다[32].

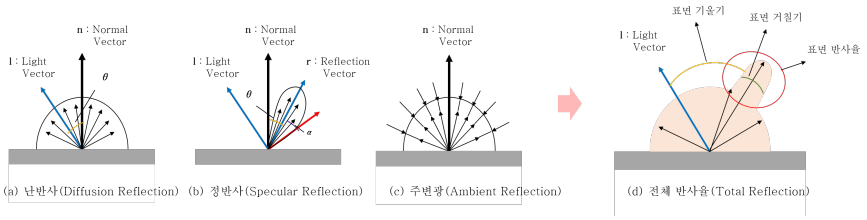


그림 27: 반사의 형태 및 방향과 전체 반사광[32]

이를 수학적으로 모델링 한 것이 Phong 모델[31]이며 식 (4.1)과 같다. 비전 센서를 통해 들어오는 물체에 대한 빛의 반사는 전체 반사광의 세기인 Phong 모델에서의 I 의 정보이다. 여기에서 K_a는 주변광 계수, I_a는 광원에서 물체면에 입사되는 입사광의 세기, $\hat{N}$은 표면 법선 벡터, $\hat{L}$은 광원벡터, K_d는 확산광 계수, K_s는 경면광 계수, I_s는 광원에서 물체면에 입사되는 입사광의 세기이며 n은 광택 계수이다.

$$
\begin{aligned}
I &= AmbientReflection + DiffuseReflection + SpecularReflection \\
&= K_a I_a + K_d I_d (\hat{N} \cdot \hat{L}) + K_s I_s (\hat{R} \cdot \hat{V})^n
\end{aligned} \tag{4.1}
$$

전체 반사광은 복합적인 정보이며 입사각에 따라, 조명의 형태에 따라 달라진다. 따라서 원하는 정보, 본 연구에서는 [그림 27]의 돌출된 영역인 표면 정보만을 얻어야한다. 하지만 이를 위해서는 전 방향에서의 반사 정보를 취합하여 후처리 및 분석하는 방법이 있을 수 있으나 시간과 비용적으로 현실성이 떨어진다. 또한 조명의 형태는 대상물, 주변 환경, 보고자 하는 목적에 따라 종류가 다양하므로 조명에 따라 얻을 수 있는 정보가 달라진다. 대부분의 조명 시스템과 현장 상황은 조명 외 주변의 빛을 완벽하게 차단하기 어려우며 평면 형태이므로 차단되지 않은 부분의 광학 데이터의 정확도가 떨어질 수 있다. 또한 낮은 각도로 조사하는 조명을 확보하기 난해하여 암시야(Dark field) 반사광을 제대로 획득할 수 없게 되므로 표면 상태의 파악이 어렵게 된다.

본 연구에서는 이러한 문제점들을 해결하기 위해 하나의 조명 시스템에서 다양한 형태의 조사광을 생성하여 다양한 이미지 정보들을 획득한 후 이를 처리하여 표면 상태의 정보만을 얻을 수 있는 비전 센서를 사용하였으며 [그림 28] 과 같이 시스템이 구성된다[33]. 비전 센서는 [그림 29]와 같이 돔(Dome)조명과 동축(Coaxial) 조명이 동시에 적용될 수 있으며 개별 조절 가능한 690개의 LED 가 존재한다 [34]. 이 다수의 광원들은 물체를 둘러싸도록 배치되어 있어 다수의 방향에서 조사가 가능하다. 센서 내 위치한 수광부에서 조사된 입사광들에 따른 물체의 반사광을 획득하

며 개별 제어 가능한 LED 광원은 미리 설정한 링(Ring Type), 플랫(Flat Type) 등의 패턴에 따라 점등되어 수광부에서 획득한 반사광들로부터 방향 별 강도 분포(Intensity Distribution)를 획득하고 이 강도분포로부터 표면의 기울기, 거칠기, 반사율을 측정하여 10ms 내 7개의 이미지를 출력한다[33]. [그림 30] 은 비전 센서로부터 얻은 7개의 이미지이며 반사율 데이터는 변색 등의 불량 검출, 거칠기의 경우 스크래치에 적합하며 기울기는 찍힘과 돌출 등의 변화를 검출할 수 있다. 본 연구에서는 표면의 기울기 이미지를 데이터로서 활용하였다.

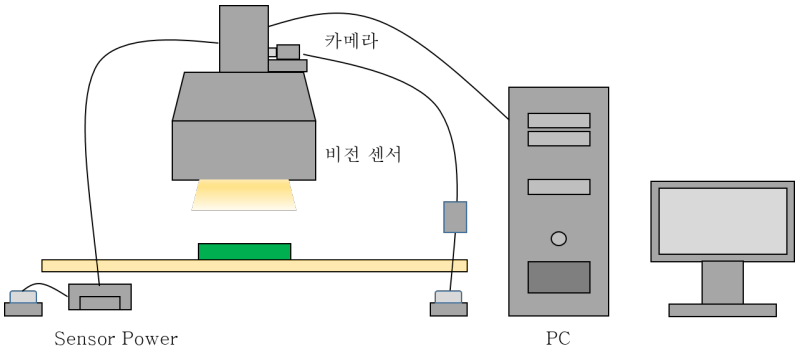
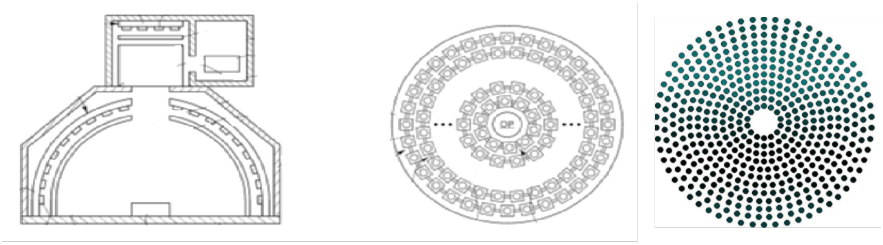


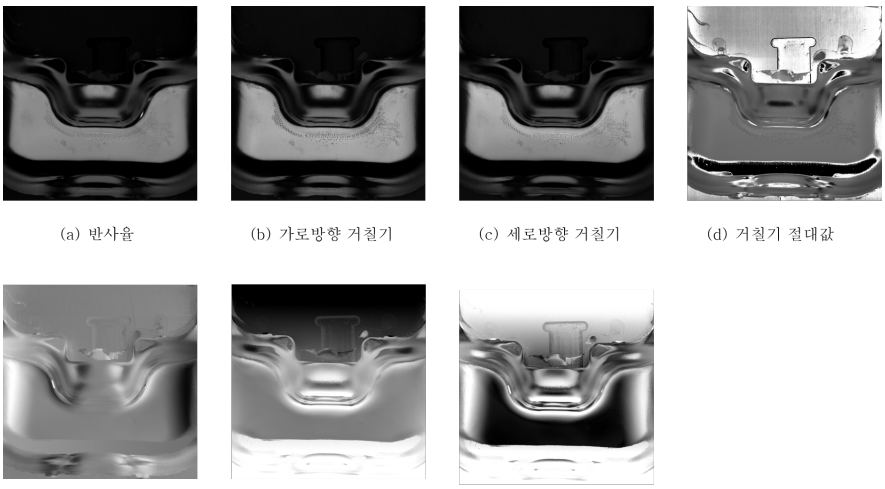
그림 28: 비전 센서 시스템 구성

본 연구의 대상 제품의 이미지 취득 실험 조건은 [그림 31]과 같이 측정 거리(Working distance)는 150mm 이며 FOV(Field of View)는 90x90mm 이다. 외부 조명은 형광등이 켜져 있는 상태에서 실험하였으며 앞서 배경에서 설명한 바와 같이 제품과 불량의 형태, 그리고 비전 센서의 FOV가 제품 사이즈인 374.4x112.8mm 보다 작음을 고려하여 Front와 Back을 나누어서 측정하였으며 카메라는 CCD카메라로 4M화소의 카메라를 사용하였다. 실험 결과 [그림 32] 와 같이 다양한 형태와 미세한 결함의 이미지



(a) 내부 단면도 (b) LED 배열

그림 29: 비전 센서 내부 단면도 및 LED 배열[34]



(a) 반사율 (b) 가로방향 거칠기 (c) 세로방향 거칠기 (d) 거칠기 절대값

(e) 가로방향 기울기 (f) 세로방향 기울기 (g) 기울기 절대값

그림 30: 비전센서로부터 얻은 7가지 데이터

도 측정이 가능한 것을 확인하였다. 다양한 종류의 단일 조명을 적용하여 촬영 시 [그림 33]과 같이 조명에 따라 결함이 정확히 표현되지 않는 경우가 발생하나 가변 조명과 원하는 형태로의 이미지 변환 시 주변 조명 환경과 무관하게 광택 표면의 미세한 결함의 측정이 가능한 것을 확인하였다. 이 실험 결과를 바탕으로 현장에 비전 센서를 설치하였으며 나누어서 측정 시 검사 속도에 문제가 발생하므로 현장에는 [그림 34] 와 같이 Front

와 Back을 동시에 검사할 수 있도록 2대의 비전 센서를 설치하였다. 이를 통해 실제 이미지 취득을 하였으며 Front 의 양품 이미지 1,352장, 불량 이미지 153장, Back의 양품 이미지 2,497장, 불량 이미지 1,374장을 얻었다.

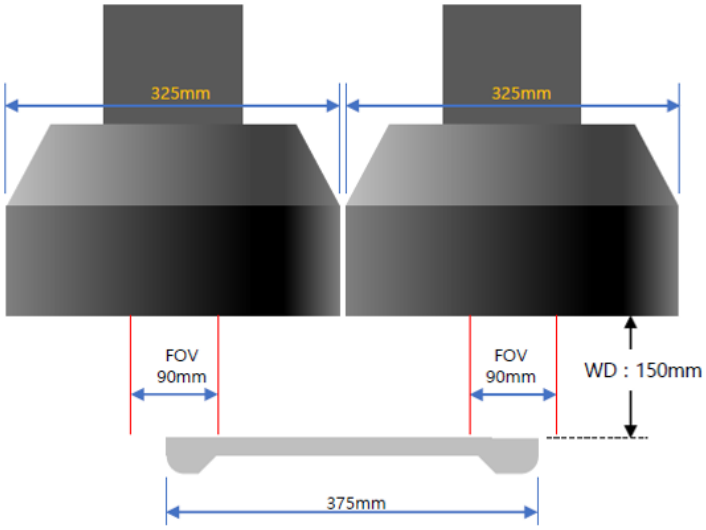
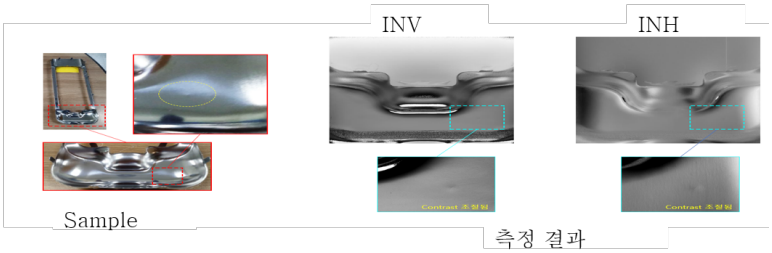
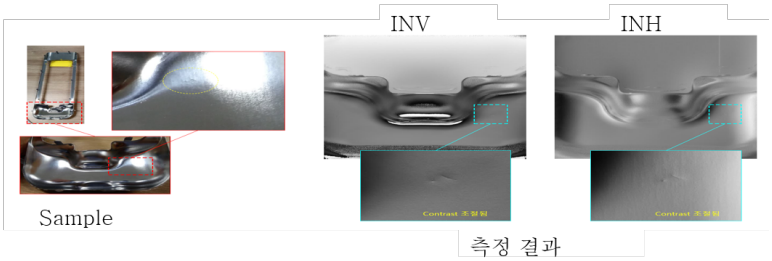


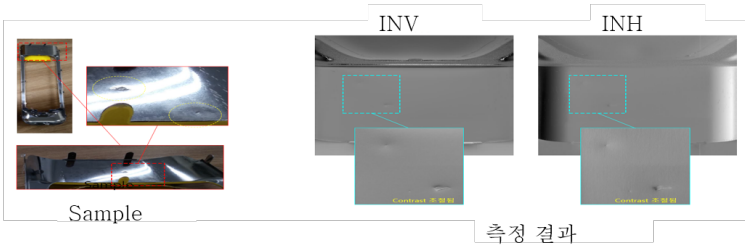
그림 31: 실험 조건 세팅



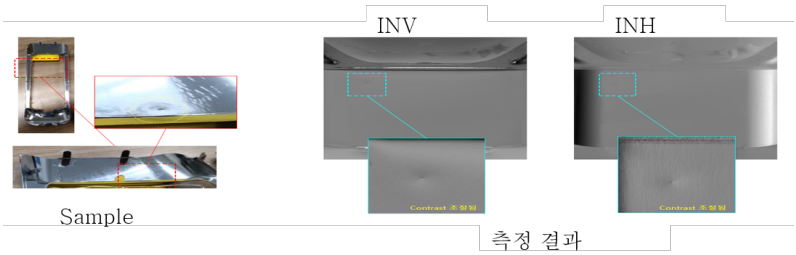
(a) Front(1)



(b) Front(2)

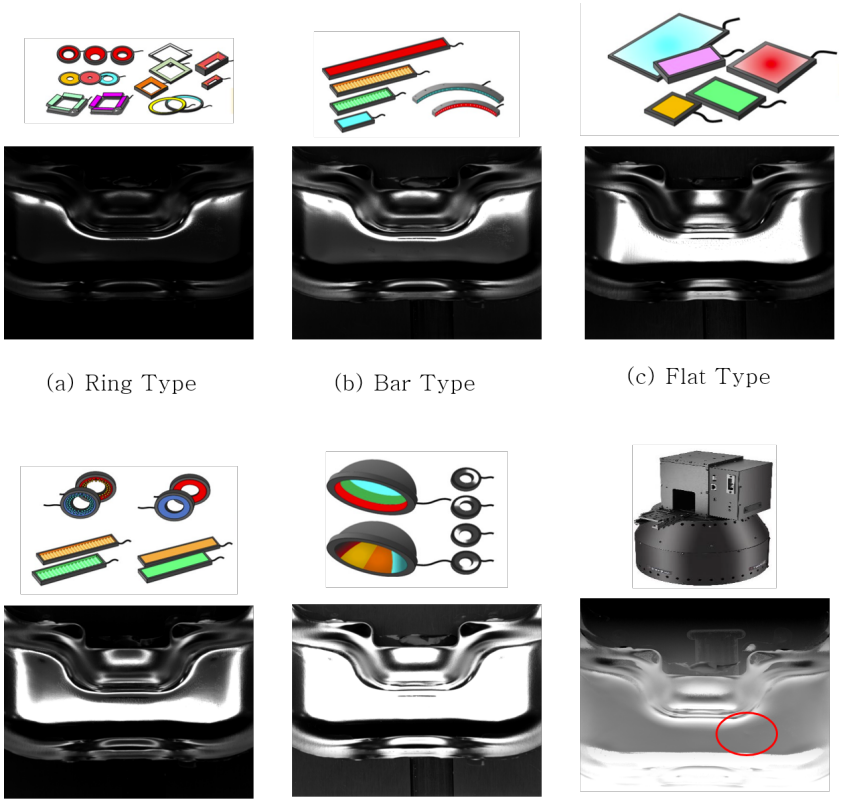


(c) Back(1)



(d) Back(2)

그림 32: Front, Back 표면 찍힘 검출 테스트



(a) Ring Type (b) Bar Type (c) Flat Type

(d) Diffusion Type (e) Dome Type (f) Multi Type

그림 33: 조명의 종류 별 취득 이미지

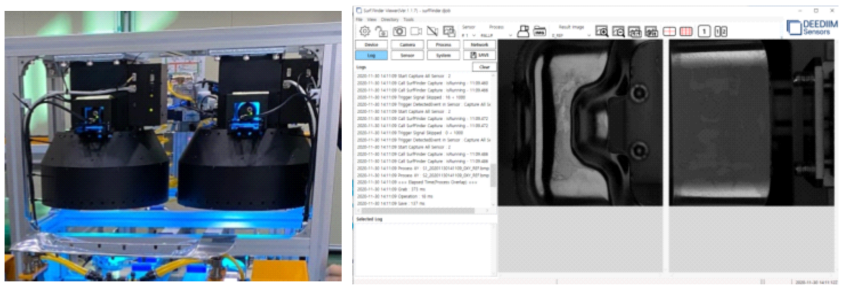
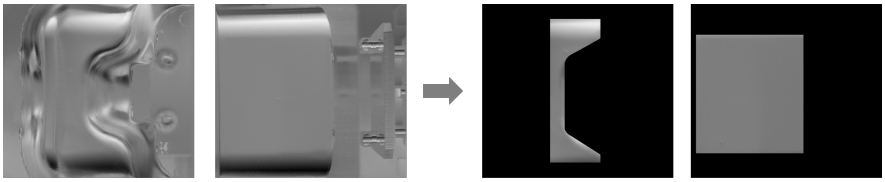


그림 34: 비전 센서 현장 설치

4.2 데이터셋 구성 및 전처리

앞서 설명하였듯이 대상 제품의 형상, 크기와 불량 유형에 따라 Front와 Back 두 가지 종류의 이미지 데이터를 취득하였다. 하지만 이미지 취득 시 제품의 위치가 일정하지 않으며 불량은 평면부에 발생하므로 굴곡부와 측면 이미지는 불필요하다. 이에 동일한 데이터셋 구성을 위하여 [그림 35]와 같이 마스킹(Masking)을 하였으며 마스킹 후 필요하지 않은 검은색 영역은 제거 하였다.



(a) Front 마스킹 전 (b) Back 마스킹 전 (c) Front 마스킹 후 (d) Back 마스킹 후

그림 35: 이미지 마스킹 전 후 비교

제거 방식은 OpenCV의 Canny Edge Detecting[35] 방식으로 잘라(Crop)내었다. Canny Edge Detection은 가우시안(Gaussian) 필터를 통해 이미지 상에서 노이즈를 제거하여 부드러운 이미지를 얻은 후 픽셀들의 기울기의 크기를 구한다. 이 후 경계선 검출기에서 최대값 만을 취하고 나머지는 무시한 후 두 개의 임계값(Double Threshold)를 통해 실제 엣지와 노이즈를 구별한 뒤 엣지의 연결성을 파악하여 검출하는 방식이다. [그림 36]은 엣지 검출을 통해 자른 최종 이미지 데이터이다. 앞 절에서 얻은 총 5,376장의 이미지를 위의 방식으로 전처리 하였다.

(a) Front Crop (b) Back Crop

그림 36: 엣지 검출을 통한 이미지 전처리

4.3 이미지 합성 및 데이터 증량

비전 센서를 통해 Front 의 양품 이미지 1,352장, 불량 이미지 153장, Back의 양품 이미지 2,497장, 불량 이미지 1,374장을 얻었으나 학습하기에는 충분하지 않은 수량이며 특히 Front의 불량 이미지의 수가 현저하게 부족하다. 실제 합성곱 신경망을 설계하여 실험한 결과 Back의 경우 정확도가 57.7% 수준이었으며 Front는 학습이 되지 않았다. 이 때의 실험 조건은 다음과 같이 설계하였다. 데이터셋은 학습과 검증을 8:2 의 비율로 나누었고 데이터 전처리는 정규화(Normalization)를 하였다.

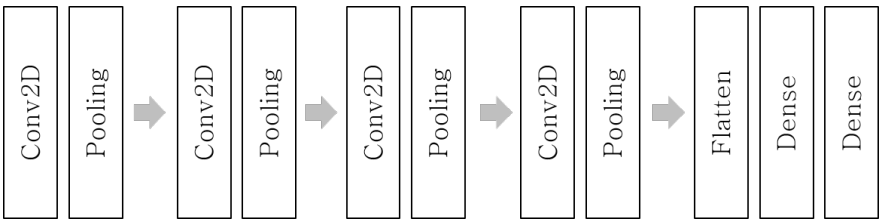
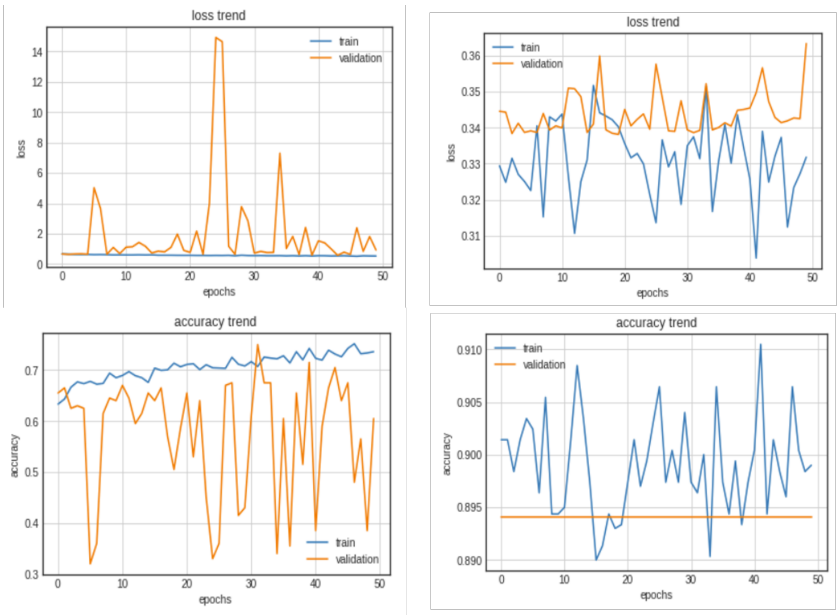


그림 37: 초기 실험 합성곱 신경망 구조

네트워크 구조는 [그림 37]과 같이 4번의 컨볼루션과 최대값 풀링, 두 번의 전연결층으로 구성하였다. 활성화 함수는 ReLu(Rectified Linear Unit)를 사용하였으며 손실함수로는 이진 교차 엔트로피(Binary Crossentropy), 최적화는 RMSprop, 학습률은 0.0005를 적용하였다. 해당 모델의 성능은 [그림 38]과 같으며 학습은 50 Epoch까지 진행한 결과이다. Front와 Back 모두 Loss와 Accuracy가 불안정하며 제대로 된 학습과 성능이 보이지 않은 것을 알 수 있었다. 이에 대한 가장 큰 원인은 데이터의 부족이라고 추정된다. 이에 데이터 증량을 진행하였으며 한정된 데이터 수량를 효율적으로 늘리기 위해 이미지 합성과 데이터 증량 분야에서 많이 활용하는 회전 및 대칭을 적용하였다.



(a) Back 영역 Loss(상), Accuracy(하)

(b) Front 영역 Loss(상), Accuracy(하)

그림 38: 초기 실험 합성곱 신경망 성능 그래프

[그림 39]는 포아송 이미지 편집 기법을 적용하는 과정이다. 포아송 이미지 편집을 위해서는 소스이미지(Source image), 마스킹(Masking), 타겟이미지(Target image)의 세 가지 이미지가 필요하다. 소스이미지는 불량 이미지로부터 추출한 결함 이미지, 마스킹은 소스 이미지의 배경이 합성되지 않도록 하는 이미지, 타겟 이미지는 추출한 결함 이미지를 합성할 이미지이다. 포아송 이미지 편집은 영역의 외곽선의 색상과 타겟 이미지가 일치되도록 픽셀간의 그래디언트가 최소화 하는 최적값을 찾는 것이며 [그림 39]의)(d)와 같이 합성하여 나온 결과물을 데이터셋에 추가하였다.

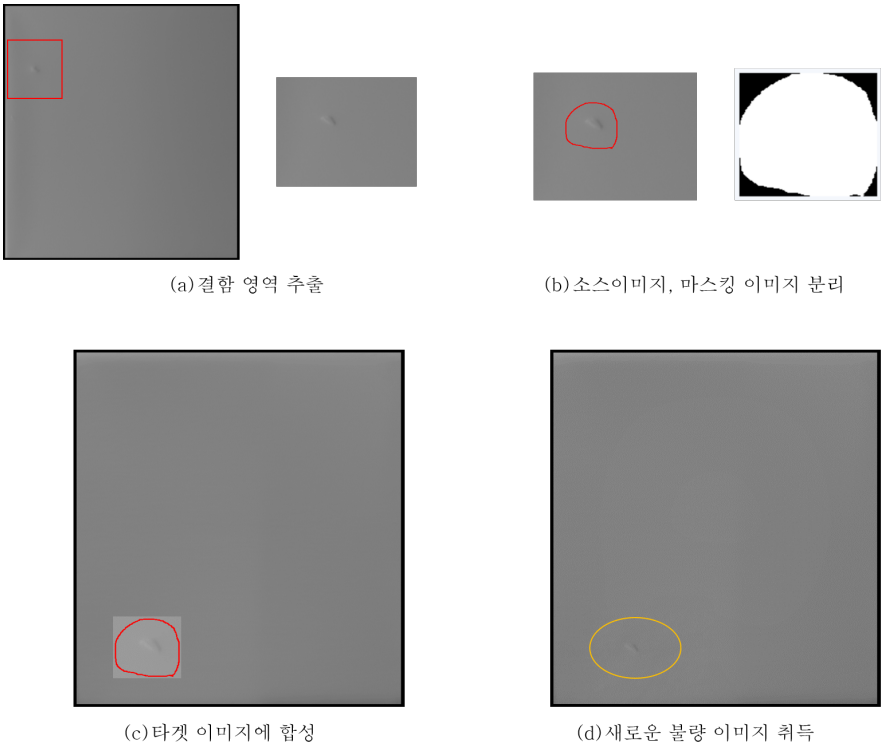
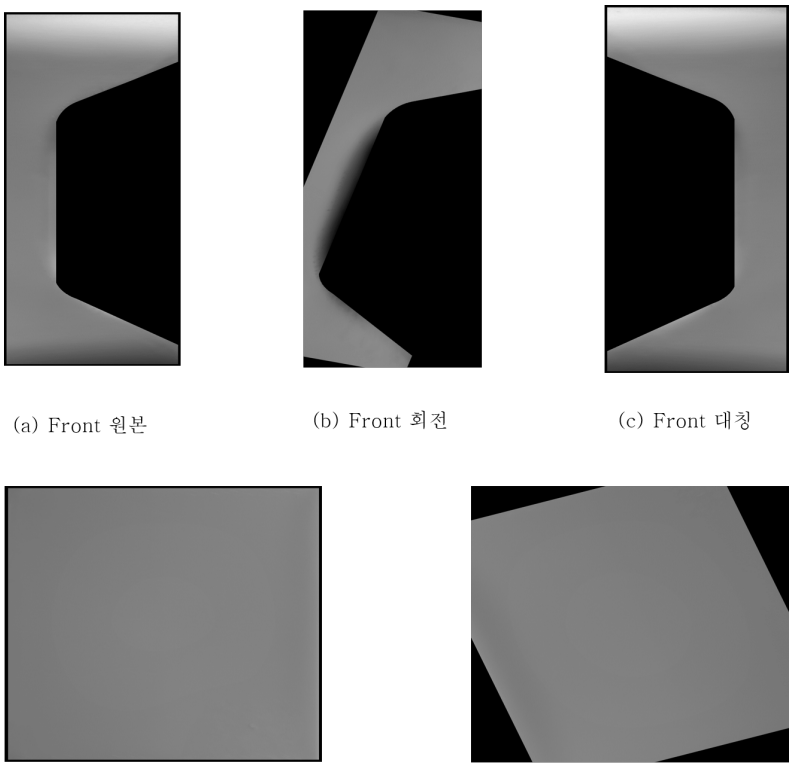


그림 39: 포아송 이미지 편집을 통한 합성 과정

이미지 합성을 통해 Front 의 불량 이미지를 기존 153장에서 1,011장으로 늘렸으며 Back의 불량 이미지를 기존 1,374장에서 2,470장으로 늘렸다. 이미지 합성을 통해 추가로 확보된 데이터셋으로 앞에서와 같은 조건으로 학습 및 검증하였을 때 Front는 70.6%의 정확도를, Back은 63.9%의 정확도를 보이며 데이터를 늘리기 전보다 개선 된 결과를 얻을 수 있었다. 하지만 여전히 낮은 성능을 보이므로 데이터 증량법을 통하여 추가 증량하였다. 데이터 증량법은 이미지 처리 분야에서 많이 사용하는 방법으로 회전, 대칭, 확대, 축소, 노이즈 등 다양한 방법을 통해 데이터를 늘리는 방법이다. 본 연구에서는 결함의 크기가 작아 데이터의 특성이 왜곡되는 것을 막기 위하여 확대, 축소 및 노이즈 추가는 적용하지 않고 [그림 40]과 같이 회전과 대칭을 통해 데이터를 증량하였다. 데이터 증량을 통해 최종 데이터셋은 Front의 양품 이미지 5,387장, Front 불량 이미지 1,839장, Back 양품이미지 7,995장, Back 불량 이미지 3,465장으로 구성되었으며 이를 정리하면 [표 2]와 같다.



(a) Front 원본 (b) Front 회전 (c) Front 대칭

(d) Back 원본 (d) Back 회전

그림 40: 회전과 대칭을 통한 데이터 증량

표 2: 데이터 증량 전,후 수량 비교

구분	Front		Back	
	양품	불량	양품	불량
증량 전(ea)	1,352	153	2,497	1,374
증량 후(ea)	5,387	1,893	7,995	3,465

4.4 실험 환경 설정

미세 결함 검출 성능 실험은 [표 3] 과 같은 환경에서 검증하였다. Ubuntu 18.04.4 OS 환경하에서 CPU는 Intel사의 i7-9700K, GPU는 NVIDIA 사의 GeForce RTX 2080Ti를 사용하였다. 사용 언어는 파이썬(Python) 3.9.7을 기반으로 텐서플로우(Tensorflow) 2.4.1 라이브러리 등을 적용하여 딥러닝 기반 이미지 분류 성능을 검증하였다. 본 연구에서는 앞서 설명한 바와 같이 기본적으로 Front와 Back 을 나누어 [표 2]의 데이터셋을 사용하여 고광택 표면의 미세 결함 탐지 분류의 성능을 검증하기 위해 다음의 실험을 진행하였다. 첫번째 실험은 전이학습을 통한 아키텍처 별 성능 비교, 두번째 실험은 학습률, 배치 크기(Batch Size), 반복(Epoch)수에 따른 과적합 개선 및 정확도 향상에 대한 검증, 마지막으로 성능이 가장 좋았던 아키텍처 2개를 선정하여 추가 성능 향상을 위한 방안에 대해 검증하였다.

표 3: 실험 환경 설정

항목	사양
OS	Ubuntu 18.04.4
GPU	NVIDIA社 GeForce RTX 2080Ti
CPU	Intel社 Core i7-9700K
Memory	64GB
Language	Python 3.9.7
Framework	Tensorflow 2.4.1

4.5 실험 결과 및 평가

4.5.1 아키텍처 별 전이학습 성능 비교

초기 소량의 데이터셋으로 합성곱 신경망을 통한 분류 성능 확인 시 Front는 제대로 학습이 되지 않았으며 상대적으로 데이터의 수가 많은 Back은 정확도가 57.5% 수준이었다. 이에 이미지 합성을 통해 데이터 수를 늘렸으며 이 때의 정확도는 Front 70.6%, Back은 64.9%를 나타내었다. 추가로 회전과 대칭을 통한 데이터 증량 후 평가 결과 Front의 경우 78.8%의 정확도를, Back은 82.3%의 정확도를 보여 데이터증량에 따른 성능 향상을 확인할 수 있었다. 하지만 여전히 낮은 정확도를 보여 실제 사용하기에는 어려운 수준이며 데이터 증량에는 한계가 있어 데이터 수량이 충분하지 않을 경우 대량의 데이터로부터 사전 학습된 가중치를 적용하여 성능을 향상 시키는 방법인 전이 학습을 통해 분류 성능을 향상 시키고자 하였다.

전이학습은 크게 세가지 단계로 진행된다. 첫번째는 사전학습 모델의 선정이다. 딥러닝 모델로써 성능이 뛰어나며 문제에 적합한 모델을 선정하는 것이다. 본 연구에서는 Top-1 정확도가 높으며 사전학습모델로써 많이 사용되는 5가지 모델인 ResNet, VGGNet, Inception V3, Xception, EfficientNet 을 선정 및 비교하였다.

본 연구에서의 전이학습 실험은 파이썬으로 작성된 오픈소스신경망 라이브러리인 케라스(Keras)를 사용하여 수행하였으며 사전 모델은 케라스 어플리케이션에서 지원하는 모델 중 앞서 말한 바와 같이 5가지를 선정하였다. 해당 모델들의 요약은 [표 4]와 같다.[36]

표 4: 전이학습을 위한 사전학습모델 비교표

Model	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth	Time (ms) per inference step (CPU)	Time (ms) per inference step (GPU)
Xception	79.0%	94.5%	2290만	81	109.4	8.1
VGG16	71.3%	90.1%	1억 3840만	16	69.5	4.2
ResNet 50	74.9%	92.1%	2560만	107	58.2	4.6
Inception V3	77.9%	93.7%	2390만	189	42.2	6.9
EfficientNet B7	84.3%	97.0%	6670만	438	1578.9	61.6

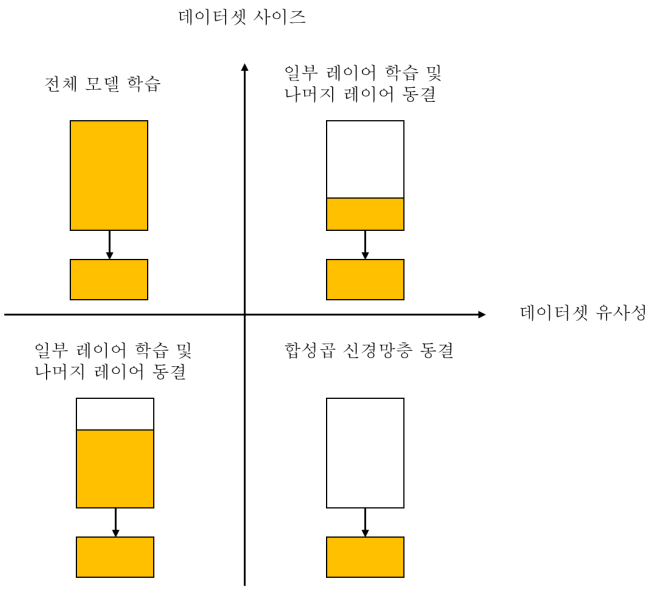


그림 41: 사전훈련모델 결정맵[37]

다음은 학습 방법의 결정이다. 현재 보유하고 있는 데이터의 수와 문제에 따라 전체 모델을 학습시킬 것인지, 분류기만 학습 시킬 것인지, 일부 레이어까지 포함하여 학습시킬 것인지 결정해야 하며 이는 [그림 41]과 같이 데이터의 수와 문제와의 유사성에 따라 달라진다.[37] 본 연구에서는 데이터수량이 충분하지 않으나 복잡하지 않은 분류 모델이므로 사전 모델이 특징 추출을 하도록 하고 분류기를 학습하는 방식을 적용하였다.

마지막은 학습 결과에 따른 미세조정이다. 앞선 과정에서 기본 모델을 구축한 후 학습 결과에 따라 기본 모델 혹은 일부를 동결 해제 하거나 학습률 조정을 통해 성능향상과 과적합을 방지하는 과정을 진행한다.

Front와 Back의 실험 설정은 동일하게 하였다. 데이터 전처리는 정규화(Normalization)를 진행하였으며 활성화 함수는 ReLu(Rectified Linear Unit)를 사용하였다. 그래디언트 폭발 혹은 소실의 방지를 위하여 합성곱층과 ReLu 사이에 Batch Normalize를 적용하였으며 [22] Loss는 Binary Crossentropy를, 최적화는 Adam(Adaptive Moment Estimation)을 사용하였다. Batch Size는 16, 학습률은 0.00002를 적용, Epoch는 50으로 설정하였으며 정리하면 [표 5]와 같다.

표 5: 전이학습 조건 요약

항목	적용
데이터 전처리	정규화
활성화 함수	ReLu
안정화	Batch Normalization
Loss	Binary Crossentropy
Optimizer	Adam
Batch Size	16
학습률	0.00002
Epochs	50

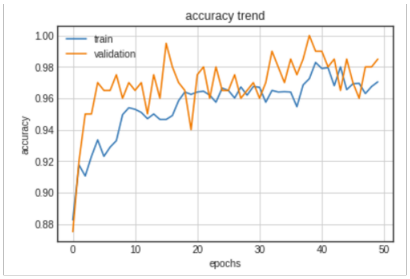
이 때의 성능을 Front와 Back, 그리고 앞서 설정한 초기 합성곱 신경망과 5개의 사전학습모델을 활용한 전이학습의 결과를 정리하면 [표 6]과 같다. 전반적으로 초기의 합성곱신경망 모델 대비 정확도가 눈에 띄게 높아졌으며 특히 Back의 경우 Inception V3와 Xception 모델에서 좋은 성능을 보이는 것을 확인하였다. 하지만 [그림 42]와 같이 Loss가 안정적

으로 수렴하지 못하며 학습과 검증의 정확도 그래프의 차이가 벌어지는 과적합(Overfitting)이 발생하였다. 또한 Front의 경우 Xception 사전학습 모델 적용 시 89.9%의 정확도를 보였으나 이는 목표로 설정한 정확도에 미치지 못한 결과이다.

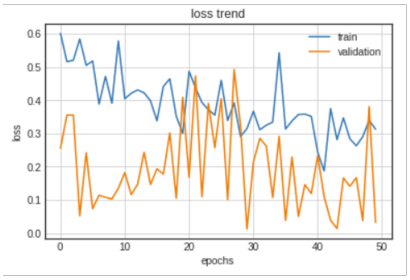
표 6: 전이학습 정확도 비교

구분	초기	전이학습				
	합성곱신경망	VGG16	Inception V3	Xception	ResNet 50	EfficientNet B7
Front	78.8%	81.3%	86.9%	89.9%	78.4%	75.6%
Back	82.3%	83.8%	96.9%	97.4%	75.0%	59.1%

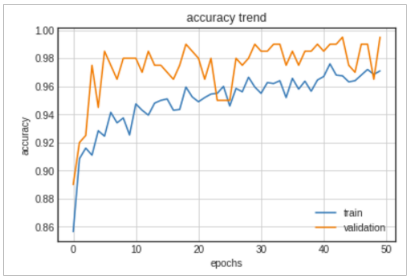
(a) Back의 Inception V3 사전모델 적용 시 Loss



(b) Back의 Inception V3 사전모델 적용 시 Accuracy



(c) Back의 Xception 사전모델 적용 시 Loss



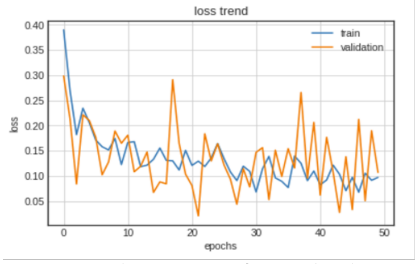
(d) Back의 Xception 사전모델 적용 시 Accuracy

그림 42: Back의 Inception V3와 Xception 전이학습 성능 그래프

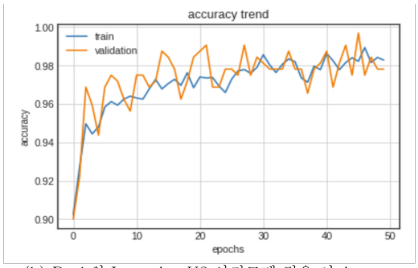
전이학습의 경우 모델의 크기 대비 데이터의 수가 적은 경우가 많으므로 학습률이 낮을수록 과적합을 방지하면서 새로운 데이터셋에 대한 모델의 성능이 향상 될 가능성이 있다[38]. 따라서 앞의 실험 조건에서 학습률을 기존 0.00002에서 0.00001로 변경하고 Batch Size는 기존 16에서 32로 변경하여 실험을 진행하였다. 그 결과는 [표 7]과 [그림 43]과 같다. 앞의 실험 결과 대비 로스가 더욱 안정적으로 감소하며 정확도는 Back의 경우 소폭 상승하였으며 Front는 최대 3.8% 향상된 결과를 나타내었다. 또한 과적합의 경향도 개선되는 것을 확인할 수 있었다. 이러한 결과를 바탕으로 학습률과 Batch Size에 따라 성능이 달라질 수 있을 것이라는 가정을 세우고 추가 실험을 진행하였다.

표 7: 학습률과 Batch Size 변경 후 전이학습 정확도 비교

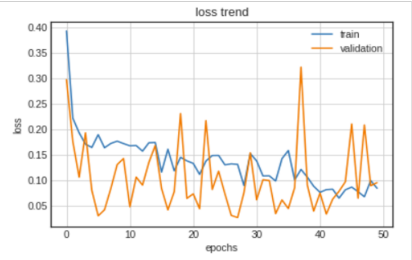
구분	초기	전이학습				
	합성곱신경망	VGG16	Inception V3	Xception	ResNet 50	EfficientNet B7
Front	78.8%	81.1%	90.7%	90.5%	78.8%	73.6%
Back	82.3%	84.3%	97.5%	97.7%	75.5%	66.8%



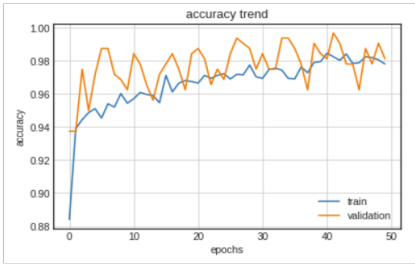
(a) Back의 Inception V3 사전모델 적용 시 Loss



(b) Back의 Inception V3 사전모델 적용 시 Accuracy



(c) Back의 Xception 사전모델 적용 시 Loss



(d) Back의 Xception 사전모델 적용 시 Accuracy

그림 43: Back 영역의 실험 설정 변경 후 전이학습 성능 그래프

4.5.2 학습률, Batch Size 별 성능 분석 - Front

4.5.1의 실험을 바탕으로 5개의 사전학습모델 중 성능이 가장 좋았던 Inception V3와 Xception에 대해 학습률, Batch Size, Epochs 의 변화에 따른 전이학습 성능을 비교하였다. 실험 조건은 앞선 실험과 같이 Front와 Back을 나누었으며 학습률은 0.001, 0.0001, 0.00002, 0.00001의 4조건, Batch Size는 4, 8, 6의 3조건, Epochs는 50, 100 2가지의 조건을 미니탭(Minitab)을 통한 실험계획법(DOE)[39]으로 수립하였다. 실험계획은 완전요인배치 법(Full Factorial Design)으로 설정하였으며 [표 8]과 같이 총 24가지의 조건이 도출되었다.

표 8: 학습률, Batch Size, Epoch 조건 별 실험 계획

Condition	Learning Rate	Epoch	Batch Size
Condition1	1.E-05	50	4
Condition2	1.E-05	100	4
Condition3	1.E-05	50	8
Condition4	1.E-05	100	8
Condition5	1.E-05	50	16
Condition6	1.E-05	100	16
Condition7	2.E-05	50	4
Condition8	2.E-05	100	4
Condition9	2.E-05	50	8
Condition10	2.E-05	100	8
Condition11	2.E-05	50	16
Condition12	2.E-05	100	16
Condition13	1.E-04	50	4
Condition14	1.E-04	100	4
Condition15	1.E-04	50	8
Condition16	1.E-04	100	8
Condition17	1.E-04	50	16
Condition18	1.E-04	100	16
Condition19	1.E-03	50	4
Condition20	1.E-03	100	4
Condition21	1.E-03	50	8
Condition22	1.E-03	100	8
Condition23	1.E-03	50	16
Condition24	1.E-03	100	16

학습률, Batch Size, Epoch 가 변수이며 이 외의 실험 조건은 [표 5]와 동일하게 설정하였다. 해당 실험에서의 정량적인 성능 지표는 정확도, 재현율, 정밀도이며 정성적으로는 과적합 여부로 성능을 판단하였다. [표 9]는 Front의 Inception V3 사전학습모델을 사용한 전이학습 성능이며 [표 10]은 Front 의 Xception 의 성능 결과이다. Inception V3와 Xcpetion을 사전학습모델로 사용하여 전이학습을 한 결과 정확도 측면에서는 학습률이 낮을수록 높은 정확도를 보였으나 Loss가 수렴하지 못하고 다시 상승하는 과적합이 발생하였다. [그림 44]는 Front영역 Inception V3의 정확도가 가장 높았던 Condition 12인 학습률 0.00002, Batch Size 16, Epochs 100

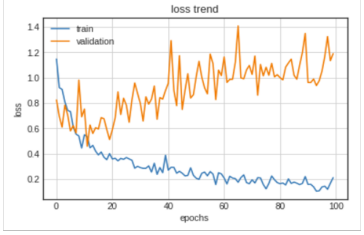
일때의 결과와 Xception의 Condition 11인 학습률 0.00002, Batch Size 16, Epochs 50 조건에서의 Loss와 정확도 그래프이다. 또한 과적합이 발생하지 않은 0.001의 학습률에서 Inception V3의 경우 Batch Size에 따른 정확도, 재현율과 정밀도에는 유의차가 없었으나 Xception의 경우에는 0.001의 학습률에서 Batch Size와 Epoch이 클수록 좋은 성능을 보였다. 결과적으로 Front 데이터셋의 경우 과적합과 성능 측면에서는 학습률과 Batch Size가 클수록 좋은 성능을 보인다고 할 수 있겠다. [그림 45]는 과적합이 없는 조건에서 정확도가 가장 높은 Inception V3의 Condition 24과 Xception의 Condition 16의 그래프이다. 과적합이 발생하지 않은 조건하에서 Front영역의 전이학습을 통한 검증 결과 정확도는 최대 96.3%로 본 연구의 목표인 정확도 97%이상 달성에는 미치지 못하였다. 이에 추가 개선을 위한 실험을 진행하였으며 이에 대한 내용은 4.5.4 에서 기술하고자 한다.

표 9: Front 실험 조건 별 Inception V3 전이학습 성능

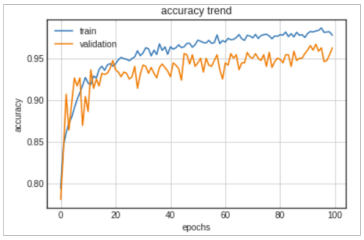
Condition	Learning Rate	Epoch	Batch Size	정확도(%)	재현율(%)	정밀도(%)	과적합 여부
Condition1	1.E-05	50	4	94.7	97.3	95.6	O
Condition2	1.E-05	100	4	95.8	97.5	96.8	O
Condition3	1.E-05	50	8	95.9	97.2	97.3	O
Condition4	1.E-05	100	8	96.2	97.7	97.2	O
Condition5	1.E-05	50	16	95.5	97	97	O
Condition6	1.E-05	100	16	95	97.2	96.1	O
Condition7	2.E-05	50	4	95.5	96.4	97.5	O
Condition8	2.E-05	100	4	94.7	97	95.9	O
Condition9	2.E-05	50	8	94	96.4	95.5	O
Condition10	2.E-05	100	8	95.8	97.3	97	O
Condition11	2.E-05	50	16	95.8	97.3	97	O
Condition12	2.E-05	100	16	96.2	97.5	97.4	O
Condition13	1.E-04	50	4	92.3	95.3	94.4	O
Condition14	1.E-04	100	4	94	94.7	97.1	O
Condition15	1.E-04	50	8	94	95.5	96.4	O
Condition16	1.E-04	100	8	94.5	97.2	95.6	O
Condition17	1.E-04	50	16	95.2	97.5	96.1	O
Condition18	1.E-04	100	16	94.1	95.5	96.6	O
Condition19	1.E-03	50	4	93.7	95.5	96	X
Condition20	1.E-03	100	4	94	97.3	94.7	X
Condition21	1.E-03	50	8	90.7	94	93.5	X
Condition22	1.E-03	100	8	94	96	95.9	X
Condition23	1.E-03	50	16	91.8	97	92.3	X
Condition24	1.E-03	100	16	94.7	97.2	95.7	X

표 10: Front 실험 조건 별 Xception 전이학습 성능

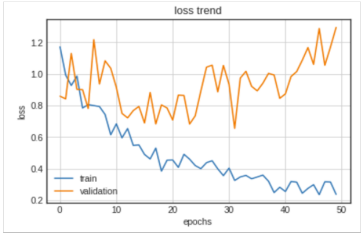
Condition	Learning Rate	Epoch	Batch Size	정확도(%)	재현율(%)	정밀도(%)	과적합 여부
Condition1	1.E-05	50	4	95.5	96	97.9	O
Condition2	1.E-05	100	4	93.3	97	94.2	O
Condition3	1.E-05	50	8	93.7	97	94.7	O
Condition4	1.E-05	100	8	95.4	96.1	97.7	O
Condition5	1.E-05	50	16	95	97.9	95.4	O
Condition6	1.E-05	100	16	95.8	97.7	96.6	O
Condition7	2.E-05	50	4	91.6	96	92.9	O
Condition8	2.E-05	100	4	95.2	96.6	97	O
Condition9	2.E-05	50	8	94.4	96.2	96.2	O
Condition10	2.E-05	100	8	95.7	97.3	96.8	O
Condition11	2.E-05	50	16	95	98.3	95.1	O
Condition12	2.E-05	100	16	95.7	97	97.2	O
Condition13	1.E-04	50	4	94.5	95.3	97.3	X
Condition14	1.E-04	100	4	94.8	96.2	96.8	X
Condition15	1.E-04	50	8	93.8	94.7	96.9	X
Condition16	1.E-04	100	8	96.3	97.7	97.4	X
Condition17	1.E-04	50	16	93.6	94.9	96.4	X
Condition18	1.E-04	100	16	94.8	95.3	97.7	X
Condition19	1.E-03	50	4	91.8	96.4	92.8	X
Condition20	1.E-03	100	4	94.5	96.4	96.2	X
Condition21	1.E-03	50	8	91.2	97.2	91.5	X
Condition22	1.E-03	100	8	94.4	97.5	95.1	X
Condition23	1.E-03	50	16	92.9	97.7	93	X
Condition24	1.E-03	100	16	94.3	96.6	95.7	X



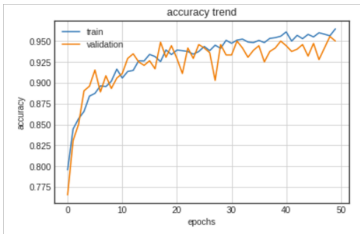
(a) Front Inception V3 Condition 12 조건의 Loss



(b) Front Inception V3 Condition 12 조건의 Accuracy

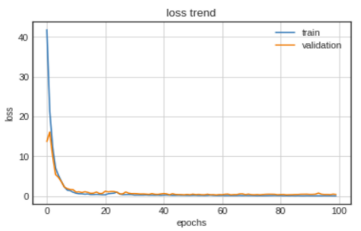


(c) Front Xception Condition 11 조건의 Loss

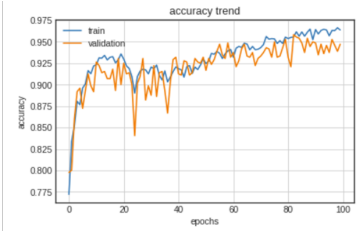


(d) Front Xception Condition 11 조건의 Accuracy

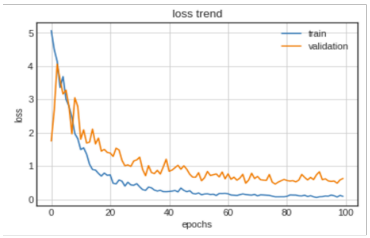
그림 44: Front Inception V3, Xception의 과적합 조건



(a) Front Inception V3 Condition 24 조건의 Loss



(b) Front Inception V3 Condition 24 조건의 Accuracy



(c) Front Xception Condition 16 조건의 Loss

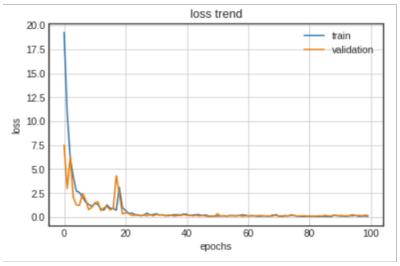
(d) Front Xception Condition 16 조건의 Accuracy

그림 45: Front Inception V3, Xception의 정상 성능 조건

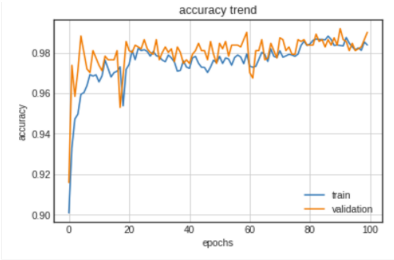
4.5.3 학습률, Batch Size 별 성능 분석 – Back

앞의 Front 영역과 같이 Back 영역에 대해서 동일한 실험을 진행하였다. 전이학습을 사용하였으며 5가지 사전학습모델 중 성능이 가장 좋았던 Inception V3, Xception 2모델에 대해 실험계획수립을 통해 조건을 설정하였다. 학습률과 Epoch는 Front와 동일하지만 [표 2]와 같이 Back의 데이터셋 크기가 Front 보다 크므로 Batch Size는 8,16,32로 설정하였다. 성능지표 또한 마찬가지로 정확도, 재현율, 정밀도이며 정성적으로 과적합 여부로 성능을 판단하였다.

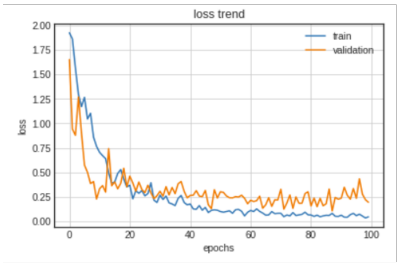
[표 11]은 Back의 Inception V3 사전학습모델을 사용한 전이학습 성능이며 [표 12]는 Back의 Xception 의 성능 결과이다. Front와 마찬가지로 Inception V3와 Xception을 사전학습모델로 하여 전이학습 한 결과 학습률이 작을수록 과적합이 발생하였다. 하지만 Front와 다른 점은 Back의 경우 학습률이 커도 정확도가 낮아지지 않았으며 오히려 더 높은 정확도를 보였다. 과적합이 발생하지 않은 학습률 0.001에서 Batch Size와 Epoch에 따른 성능 평가에는 유의차가 없었다. [그림 46]은 과적합이 없는 조건에서 정확도가 가장 높은 Inception V3의 Condition 20과 Xception의 Condition 18의 그래프이다. 과적합이 발생하지 않은 조건하에서 Back 영역의 전이학습을 통한 검증 결과 정확도는 Inception V3와 Xception 모두 99%의 정확도를 보여 본 연구의 목표인 정확도 97%이상의 성능을 확인하였다.



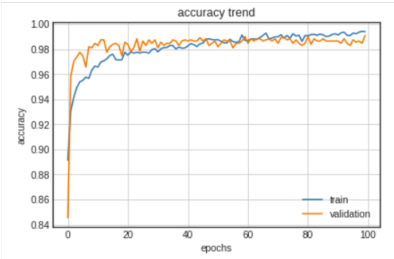
(a) Back Inception V3 Condition 20 조건의 Loss



(b) Back Inception V3 Condition 20 조건의 Accuracy



(c) Back Xception Condition 18 조건의 Loss



(d) Back Xception Condition 18 조건의 Accuracy

그림 46: Back Inception V3, Xception의 정상 성능 조건

표 11: Back 실험 조건 별 Inception V3 전이학습 성능

Condition	Learning Rate	Epoch	Batch Size	정확도(%)	재현율(%)	정밀도(%)	과적합 여부
Condition1	1.E-05	50	8	98.5	99.2	98.7	O
Condition2	1.E-05	100	8	99	99.8	98.7	O
Condition3	1.E-05	50	16	98.5	99.7	98.2	O
Condition4	1.E-05	100	16	98.4	99.3	98.5	O
Condition5	1.E-05	50	32	98.8	99.4	98.8	O
Condition6	1.E-05	100	32	98.5	99.7	98.2	O
Condition7	2.E-05	50	8	98.9	99.7	98.7	O
Condition8	2.E-05	100	8	98.4	99.3	98.5	O
Condition9	2.E-05	50	16	98.9	99.7	98.7	O
Condition10	2.E-05	100	16	98.7	99.6	98.6	O
Condition11	2.E-05	50	32	98.8	99.7	98.6	O
Condition12	2.E-05	100	32	98.5	99.6	98.3	O
Condition13	1.E-04	50	8	98.9	99.4	99	△
Condition14	1.E-04	100	8	98.3	99.7	98	O
Condition15	1.E-04	50	16	98.8	99.7	98.6	O
Condition16	1.E-04	100	16	98.6	99.3	98.7	O
Condition17	1.E-04	50	32	98.2	99.8	97.7	O
Condition18	1.E-04	100	32	98.7	99.7	95.5	O
Condition19	1.E-03	50	8	98.6	99.2	98.8	X
Condition20	1.E-03	100	8	99	99.7	98.8	X
Condition21	1.E-03	50	16	98.1	99.3	98.1	X
Condition22	1.E-03	100	16	98.4	99.7	98.1	X
Condition23	1.E-03	50	32	98.4	98.9	98.8	X
Condition24	1.E-03	100	32	98.2	99.1	98.5	X

표 12: Back 실험 조건 별 Xception 전이학습 성능

Condition	Learning Rate	Epoch	Batch Size	정확도(%)	재현율(%)	정밀도(%)	과적합 여부
Condition1	1.E-05	50	8	98.7	99.6	98.6	O
Condition2	1.E-05	100	8	98.1	99.8	97.6	O
Condition3	1.E-05	50	16	98.4	99.6	98.2	O
Condition4	1.E-05	100	16	98.1	99.7	97.7	O
Condition5	1.E-05	50	32	98.3	99.6	98.1	O
Condition6	1.E-05	100	32	98.4	99.8	98	O
Condition7	2.E-05	50	8	98.8	99.7	98.6	O
Condition8	2.E-05	100	8	98.3	99.8	97.9	O
Condition9	2.E-05	50	16	98.2	99.4	98.1	O
Condition10	2.E-05	100	16	98.8	99.7	98.6	O
Condition11	2.E-05	50	32	98.1	99.1	98.2	O
Condition12	2.E-05	100	32	98.6	99.7	98.3	O
Condition13	1.E-04	50	8	98.6	99.7	98.3	△
Condition14	1.E-04	100	8	98.6	99.7	98.3	△
Condition15	1.E-04	50	16	98.1	99.1	98.3	△
Condition16	1.E-04	100	16	98.8	99.6	98.7	△
Condition17	1.E-04	50	32	98.4	99.8	98	X
Condition18	1.E-04	100	32	99	99.8	98.8	X
Condition19	1.E-03	50	8	98.6	99.7	98.3	X
Condition20	1.E-03	100	8	99	99.6	99	X
Condition21	1.E-03	50	16	98.5	99.8	98.1	X
Condition22	1.E-03	100	16	98.8	99.7	98.6	X
Condition23	1.E-03	50	32	98.8	99.4	98.8	X
Condition24	1.E-03	100	32	98.9	99.7	98.7	X

4.5.4 이미지 분할 학습을 통한 성능 향상 검증

Back 영역은 Inception V3와 Xception을 사전학습모델로 전이학습을 통한 성능 검증 시 99%의 정확도를 보여 양품과 불량의 분류 성능이 뛰어나다는 것을 확인하였으나 Front 영역의 분류 정확도는 최대 96.3%로 실제 현장에 적용하기에는 부족한 성능을 보였다. 그 원인으로는 데이터수의 부족, 전체 이미지 내 과소한 결함의 크기, 이미지 내 불필요한 영역 존재로 판단되었다. 이에 이미지의 분할과 추가 합성을 통하여 부족한 데이터셋의 보완, 용이한 특징맵 추출, 불필요한 학습 이미지를 제거하는 방안을 강구하였다. 이를 위한 순서는 다음과 같다.

첫번째로 현 보유중인 불량이미지를 [그림 47]과 같이 분할한다.

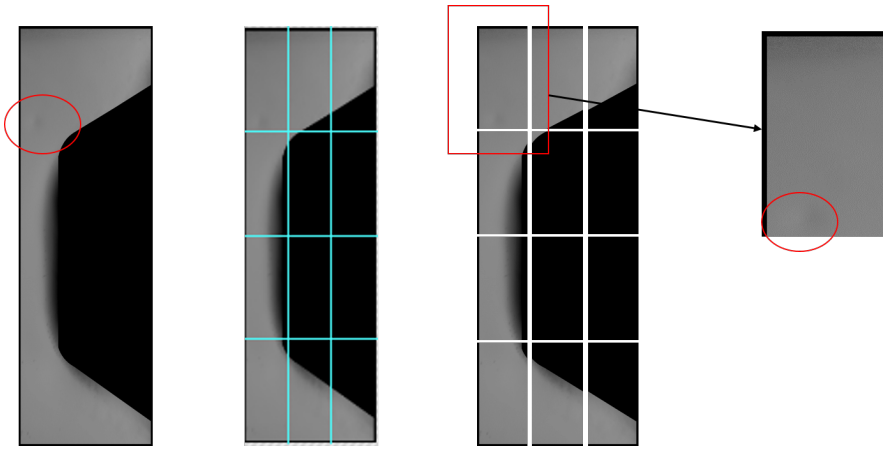


그림 47: Front 불량 이미지 분할

두번째는 이미지의 전처리이다. [그림 47]의 분할된 영역에서 제품이 아닌 검은색 영역은 불필요한 부분이다. 해당 영역은 학습과 합불 분류와 관련 없는 영역이므로 삭제할 필요가 있다. 불필요 영역 제거는 우선 육안으로 검은색 영역이 전체 이미지의 2/3 이상 차지하고 있는 이미지를 선정

후 그레이 스케일(Gray Scale) 로 변환하여 화소강도(Pixel Intensity)를 히스토그램으로 확인하였다. 히스토그램을 확인 시 검은색 영역이 2/3 이상 차지하고 있는 이미지의 경우 0부터 255 중 검은색인 0에 해당하는 픽셀의 수가 160,000이상인 것을 확인하였다. 이 후 전체 분할된 이미지에 대해서 0부터 255 중 0에 해당하는 픽셀수가 160,000이 넘어가는 이미지는 일괄 삭제하였으며 과정은 [그림 48]과 같다.

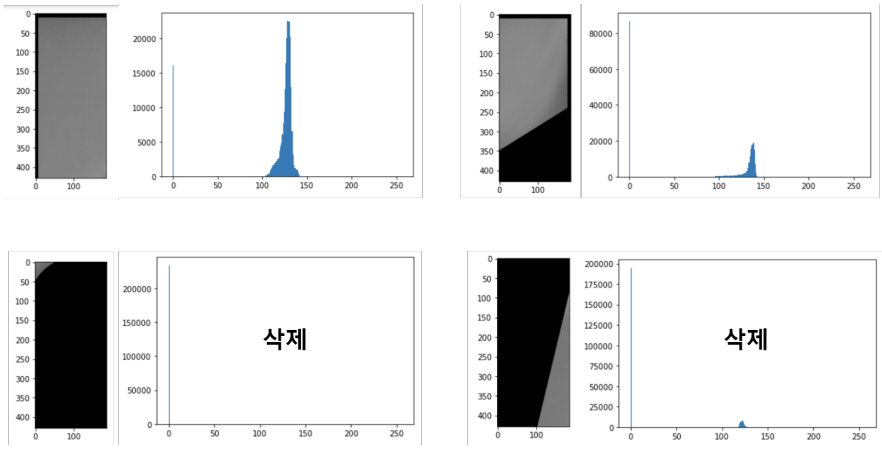
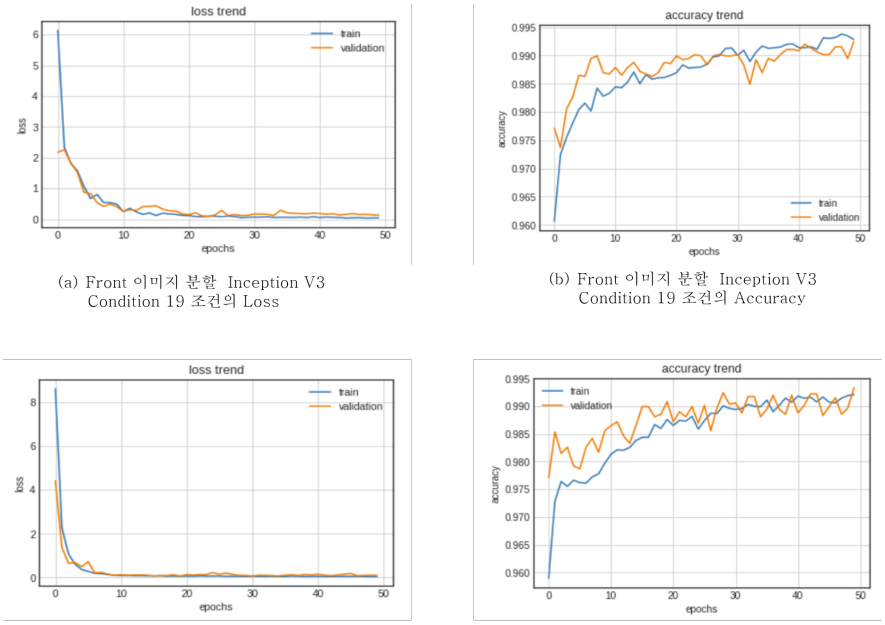


그림 48: Front 불필요 영역 삭제

세번째는 이미지의 증량이다. 앞서 데이터셋의 수량을 늘린 방법으로 포와송 이미지 편집 기법을 통한 이미지 합성, 그리고 회전과 대칭을 통하여 전체 데이터셋의 크기를 키웠다. 증량된 이미지의 수는 양품 37,651장, 불량 이미지 2,291장으로 4.5.2의 실험 대비 양품 이미지의 수량이 크게 증가하였다. 위의 과정으로 구성된 데이터셋을 사용하여 앞선 실험과 동일한 조건인 Inception V3, Xception 을 사전학습모델로 전이학습을 진행하였다. 또한 학습률, Batch Size, Epoch와 같은 실험 설정도 동일하게 진행하였다. [표 13]과 [표 14]는 이미지 분할을 통한 전이학습 시 Inception V3와 Xception의 결과이다. 정확도의 경우 모든 조건에서 99% 이상

의 성능을 보였으나 앞선 실험과 같이 작은 학습률에서는 과적합이 발생하였다. 과적합이 발생하지 않은 학습률 0.001에서 Batch Size와 Epoch과 성능에는 유의차가 없었다. [그림 49]는 Inception V3의 Condition 19와 Xception의 Condition 23의 성능 그래프이다. 이미지 분할을 통한 전이학습을 통한 검증 결과 정확도는 Inception V3와 Xception 모두 99% 이상의 정확도를 보여 본 연구의 목표인 정확도 97%이상의 성능을 달성하였으며 이미지 분할 학습을 통해 기존 성능 대비 향상된 것을 확인할 수 있었다.



(a) Front 이미지 분할 Inception V3 Condition 19 조건의 Loss

(b) Front 이미지 분할 Inception V3 Condition 19 조건의 Accuracy

(c) Front 이미지 분할 Xception Condition 23 조건의 Loss

(d) Front 이미지 분할 Xception Condition 23 조건의 Accuracy

그림 49: Front 이미지 분할 Inception V3, Xception의 성능

표 13: Front 이미지 분할 실험 조건 별 Inception V3 전이학습 성능

Condition	Learning Rate	Epoch	Batch Size	정확도(%)	재현율(%)	정밀도(%)	과적합 여부
Condition1	1.E-05	50	4	99.2	99.8	99.3	O
Condition2	1.E-05	100	4	99.2	99.8	99.3	O
Condition3	1.E-05	50	8	99.2	99.9	99.3	O
Condition4	1.E-05	100	8	99.2	99.9	99.3	O
Condition5	1.E-05	50	16	99.2	99.8	99.3	O
Condition6	1.E-05	100	16	99.3	99.9	99.4	O
Condition7	2.E-05	50	4	99.1	99.7	99.4	O
Condition8	2.E-05	100	4	99.2	99.9	99.2	O
Condition9	2.E-05	50	8	99.2	99.8	99.3	O
Condition10	2.E-05	100	8	99.3	99.9	99.3	O
Condition11	2.E-05	50	16	99.1	99.7	99.3	O
Condition12	2.E-05	100	16	99.2	99.8	99.3	O
Condition13	1.E-04	50	4	99.3	99.8	99.5	O
Condition14	1.E-04	100	4	99.3	99.8	99.4	O
Condition15	1.E-04	50	8	99.1	99.8	99.3	O
Condition16	1.E-04	100	8	99.2	99.6	99.5	O
Condition17	1.E-04	50	16	99.2	99.8	99.3	O
Condition18	1.E-04	100	16	99.2	99.9	99.2	O
Condition19	1.E-03	50	4	99.2	99.8	99.3	X
Condition20	1.E-03	100	4	99.1	99.7	99.3	X
Condition21	1.E-03	50	8	99	99.7	99.1	X
Condition22	1.E-03	100	8	99.1	99.8	99.3	X
Condition23	1.E-03	50	16	99	99.8	99.1	X
Condition24	1.E-03	100	16	99	99.7	99.3	X

표 14: Front 이미지 분할 실험 조건 별 Xception 전이학습 성능

Condition	Learning Rate	Epoch	Batch Size	정확도(%)	재현율(%)	정밀도(%)	과적합 여부
Condition1	1.E-05	50	4	99.1	99.8	99.3	O
Condition2	1.E-05	100	4	99.1	99.8	99.3	O
Condition3	1.E-05	50	8	99.2	99.8	99.4	O
Condition4	1.E-05	100	8	99.2	99.8	99.4	O
Condition5	1.E-05	50	16	99.2	99.8	99.3	O
Condition6	1.E-05	100	16	99.3	99.9	99.3	O
Condition7	2.E-05	50	4	99.3	99.7	99.4	O
Condition8	2.E-05	100	4	99.1	99.8	99.3	O
Condition9	2.E-05	50	8	99.1	99.8	99.2	O
Condition10	2.E-05	100	8	99.2	99.8	99.4	O
Condition11	2.E-05	50	16	99.2	99.9	99.3	O
Condition12	2.E-05	100	16	99.2	99.8	99.3	O
Condition13	1.E-04	50	4	99.1	99.8	99.2	X
Condition14	1.E-04	100	4	99.3	99.9	99.3	X
Condition15	1.E-04	50	8	99.3	99.8	99.4	X
Condition16	1.E-04	100	8	95.6	100	95.6	X
Condition17	1.E-04	50	16	99.1	99.6	99.3	X
Condition18	1.E-04	100	16	99.2	99.8	99.4	X
Condition19	1.E-03	50	4	99.1	99.6	99.5	X
Condition20	1.E-03	100	4	99.2	99.6	99.5	X
Condition21	1.E-03	50	8	99.2	99.8	99.3	X
Condition22	1.E-03	100	8	99	99.3	99.6	X
Condition23	1.E-03	50	16	99.3	99.8	99.4	X
Condition24	1.E-03	100	16	99.2	99.6	99.5	X

제 5 장

결론

본 연구보고서는 고광택 표면을 가진 제품 표면의 미세한 결함을 검출하기 위한 방안을 제시하였다. 빛이 물체에 닿을 때 난반사, 정반사, 주변광 등 복합적인 반사 형태를 보임에 따라 원하는 표면 정보만을 얻기가 어렵다. 특히 고광택의 표면은 단일 조명 사용 시 과도한 빛의 반사에 따라 광학데이터의 정확도가 낮다. 이를 극복하고자 복합적인 반사정보 내에서 원하는 표면의 상태만을 얻기 위하여 하나의 조명 시스템에서 다양한 형태의 조사광을 생성할 수 있도록 비전센서 내 개별 조절 가능한 가변 조명을 적용하였다. 다수의 방향에서 미리 설정한 패턴에 따라 점등되어 수광부에서 획득한 반사광들로부터 원하는 표면을 추출하는 방식이다. 이를 통해 고광택 표면을 가진 제품의 미세한 결함을 광학데이터로써 얻을 수 있었다. 제품의 특성 상 Front와 Back으로 나누어 데이터를 획득하였으며 Front 의 양품 이미지 1,352장, 불량 이미지 153장, Back의 양품 이미지 2,497장, 불량 이미지 1,374장을 얻어 이를 학습을 위한 데이터셋으로 사용하였다.

산업 현장에서 발생하는 데이터부족 문제를 해결하기 위하여 포아송 이미지 편집을 이용한 이미지 합성과 회전, 반전의 데이터 증량 방식을 통해 Front의 양품 이미지 5,387장, Front 불량 이미지 1,839장, Back 양품 이미지 7,995장, Back 불량 이미지 3,465장으로 증량하였다.

해당 데이터셋을 통해 기본적인 합성곱 신경망을 활용하여 학습을 진행하였으나 여전히 데이터의 수가 충분하지 않아 목표 성능에는 도달하지 못했다. 이를 해결하기 위하여 낮은 데이터셋으로 높은 성능을 보일 수 있는 전이학습을 사용하였다. 전이학습을 위한 사전학습모델로는 ResNet, VGGNet, Inception V3, Xception, EfficientNet 의 5가지를 비교하였으며 이 중 Inception V3와 Xception의 성능이 가장 높았다. 또한 학습률과 Batch Size에 따른 정확도의 관계를 확인하고자 실험계획법을 통해 각각의 조건 별로 실험을 진행하였다. 이를 통해 본 연구에서 사용한 데이터셋의 경우 학습률이 작을수록 정확도는 향상되나 과적합이 발생하며 학습률이 클수록 과적합의 개선과 성능의 향상이 이루어짐을 확인하였다. 적합한 사전학습모델의 선정과 적절한 전이학습의 미세조정을 통해 Front의 분류 정확도 96.3%, Back의 분류 정확도 99%의 성능을 나타내었다. Back의 경우 현장에 적용 가능한 수준이지만 Front는 정확도가 목표에 미치지 않아 추가 성능 향상을 위한 실험을 진행하였다.

Front의 추가 성능 향상을 위하여 이미지의 분할과 추가 합성을 통하여 부족한 데이터셋의 보완, 용이한 특징맵 추출, 불필요한 학습 이미지를 제거하는 방법을 적용하였다. 불량 이미지를 분할한 후 불필요 영역은 화소 강도 확인을 통해 삭제하는 이미지 전처리를 수행하였고 분할 된 이미지를 앞선 실험과 마찬가지로 포아송 이미지 편집 기법을 통한 합성과 회전, 반전의 이미지 증량을 하였다. 이를 통해 양품 37,651장, 불량 이미지 2,291장으로 증량하였으며 앞의 실험과 같이 Inception V3, Xception의 사전학습모델을 적용하여 전이학습을 시행하였다. 학습률과 Batch Size 별 정확도 차이를 비교하였으며 앞선 실험과 동일한 관계를 확인하였다. 이

미지 분할을 통한 전이학습 시 Inception V3, Xception 사전학습모델에서 성능이 향상됨을 확인하였고 정확도 99.3%, 재현율 99.8%, 정밀도 99.4%를 달성하였다.

본 연구는 고광택 제품의 미세 결함탐지의 정확도를 높이기 위해 가변 조명의 비전 센서를 통해 높은 정확도의 광학데이터와 이를 바탕으로 딥러닝 기반 검출 모델의 성능을 향상시키는 방안을 제시하였다. 데이터가 부족한 현장, 제품의 특성과 주변 환경으로 인해 아직 자동화 검사를 적용하지 못하고 전문 검사원이 육안검사를 하는 현장에 도움이 될 것으로 판단한다. 하지만 실제 현장에서 사용 가능한 실시간 검출 속도에 대한 검증이 아직 이루어지지 않았다. 향 후 추가적인 데이터 확보, 이미지 분할과 학습 모델의 단순화와 실시간 검출 속도 검증을 통해 현장에서 실제 사용 되어 자동화 검사를 시행할 수 있도록 하고자 한다.

참고 문헌

[1] Y.-K. Son, C.-J. Lee, S.-D. Byeon, M.-D. Kim, and B.-M. Kim, "Methods for suppressing tearing of pet coating during forming of vcm steel sheet for fabricating washer," *Transactions of the Korean Society of Mechanical Engineers A*, vol. 35, no. 9, pp. 1027–1033, 2011.

[2] Y. Lim, W.-g. Cha, S. Ko, and N. Kim, "Validation of formability of laminated sheet metal for deep drawing process using gtn damage model," in *AIP Conference Proceedings*, vol. 1567, pp. 607–613, American Institute of Physics, 2013.

[3] Y. Lim, T. An, S. Ko, and N. Kim, "Formability of coated vinyl on sheet metal during deep drawing process," *Journal of Materials Processing Technology*, vol. 227, pp. 178–189, 2016.

[4] K.-M. Yeo, J.-W. Seo, S.-W. Lee, and J.-H. Yi, "A machine-vision based inspection system for non-transparent and high-reflectance substrate," in *Proceedings of the Korea Information Processing Society Conference*, pp. 369–372, Korea Information Processing Society, 2010.

[5] J. G. Keramas, "Robot technology fundamentals,© 1999," *Thompson Learning, USA*, pp. 193–219.

[6] J.-H. Lee and M.-K. Lee, "A camera calibration algorithm for an ill-conditioned case," *Journal of the Korean Society for Precision Engineering*, vol. 16, no. 2, pp. 164–175, 1999.

[7] K.-M. Huh, S.-H. Park, and S.-M. Kang, "Vision inspection method development which improves accuracy by using power-law transformation and histogram specification," *Journal of the Institute of Electronics Engineers of Korea SC*, vol. 44, no. 5, pp. 11–17, 2007.

[8] J. Kim and C. Kim, "Obtaining shapes of specular objects using ring illumination," *Journal of the Korean Society for Precision Engineering*, vol. 12, no. 4, pp. 78–87, 1995.

[9] K.-H. Song and K.-M. Huh, "Development of automatic remocon inspection system using machine vision," in *Proceedings of the KIEE Conference*, pp. 138–140, The Korean Institute of Electrical Engineers, 2004.

[10] S. Juds, *Photoelectric sensors and controls: selection and application*, vol. 63. CRC Press, 1988.

[11] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.," *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[13] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[14] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.

[15] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural computation*, vol. 1, no. 4, pp. 541–551, 1989.

[16] S. Behnke, *Hierarchical neural networks for image interpretation*, vol. 2766. Springer, 2003.

[17] P. Y. Simard, D. Steinkraus, J. C. Platt, *et al.*, "Best practices for convolutional neural networks applied to visual document analysis.," in *Icdar*, vol. 3, Edinburgh, 2003.

[18] I.-H. Kim and I.-H. Jung, "A study on korea sign language motion recognition using openpose based on deep learning," *Journal of Digital Contents Society*, vol. 22, no. 4, pp. 681–687, 2021.

[19] TAEWAN.KIM, "Cnn, convolutional neural network summary," 2018.

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[21] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[22] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, pp. 448–456, PMLR, 2015.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034, 2015.

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[26] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, 2017.

[27] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, pp. 6105–6114, PMLR, 2019.

[28] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[29] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," in *ACM SIGGRAPH 2003 Papers*, pp. 313–318, 2003.

[30] A. Said and V. Torra, "Data science: An introduction," in *Data Science in Practice*, pp. 1–6, Springer, 2019.

[31] B. T. Phong, "Illumination for computer generated pictures," *Communications of the ACM*, vol. 18, no. 6, pp. 311–317, 1975.

[32] T. Kurihara and Y. Takaki, "Shading of a computer-generated hologram by zone plate modulation," *Optics express*, vol. 20, no. 4, pp. 3529–3540, 2012.

[33] J.-I. Ser, "Illumination system," June 1 2021. US Patent 11,026,313.

[34] J.-I. Ser, "Illumination system for recognizing material and method of recognizing material using the same," Oct. 22 2019. US Patent 10,451,547.

[35] Z. Xu, X. Baojie, and W. Guoxin, "Canny edge detection based on open cv," in *2017 13th IEEE international conference on electronic measurement & instruments (ICEMI)*, pp. 53–56, IEEE, 2017.

[36] K. Team, "Keras documentation: About keras," *Keras. io*, 2020.

[37] P. Marcelino, "Transfer learning from pre-trained models how to solve any image classification problem quickly and easily,"

[38] D. Mwiti., "A practical tutorial with examples for images and text in keras," 2017.

[39] R. A. Fisher *et al.*, “The design of experiments.,” *The design of experiments.*, no. 2nd Ed, 1937.

Abstract

Improvement of Micro-Defect Detection Accuracy on High-Glossy Surface Based on Deep Learning

KO Sang Jin
Graduate School of Practical Engineering
Seoul National University

With the development of technology and the recent establishment of smart factories, various automation cases have been shown in diverse manufacturing industries. They have improved productivity and reduced labor costs through automation and are further advancing performances through deep learning technologies. In particular, with the help of deep-learning-based image recognition methods, various inspection systems are getting improved in their accuracy and efficiency. However, there are still areas where visual inspection by a professional human inspector is still required for reasons such as diverse product shapes, surface properties, and defect states, where automatic inspection reliability is not secured. In order to perform automated inspection through deep learning, a vision system that can obtain a measurable image must be built, and a sufficient amount of data is

required for learning to achieve the desired performance.

In this research report, we propose an image acquisition method and deep-learning-based vision inspection system for detecting small and atypical defects on a product with a high gloss surface. A typical problem in learning-based visual inspection is the lack of anomaly data for the training, and thus we also propose a new data augmentation method. Specifically, we employ the Poisson image editing method to augment the dataset. To alleviate the difficult light-reflection problem due to the glossy surface, we also introduce new lighting structures and image acquisition processes.

We also attempt to find the balances between the metrics, such as Accuracy, Precision, and Recall, such that our method can be effectively applied to industrial applications. To achieve this goal, transfer learning was applied to compensate for the lack of data and improve accuracy. In addition, in order to reduce the failures in feature extraction and recognition due to the small size of the defect compared to the original image, the input image was separated into an effective size for learning, and the model with the best performance was selected by comparing each pre-trained model. The extensive experiments confirmed that our method achieves an accuracy of 99.3% by deriving the optimal conditions by changing the parameters through the DOE.

Keywords : Lighting, Vision Sensor, Data Augmentation, Deep Learning, Convolutional Neural Network, Transfer Learning

Student Number : 2021-20491