

**CNN 알고리즘 기반의
카메라모듈 얼룩불량 분류 연구**

2021. 12. 9
김병근 (2020254013)

연구 개요

연구 목표

- 카메라 모듈의 “최종 검사장비(머신 비전)”와 “CNN 분류 알고리즘”의 불량 판정 정확도 비교



연구 목표 및 기여

- 현재: 1개월(2021.11월) 비전 검사 장비의 판정 정확도 측정 (48M, 4,800만화소)
불량 제품 재검사 3,346,017 中 → 2,484,252 양품(오 검) / 861,765 (실제) 불량
- 연구: CNN(resnet18) 모델로 Classification 정확도 측정

항목	단위	현재(머신 비전 장비)	연구목표(딥러닝)	비고
불량 판정 정확도 (Classification)	%	판정 정확도 95.40%	분류 정확도 96.00%	Confusion Matrix
오(誤) 판정 부대비용		재검사 용(用) 비용 발생 장비 구입/공간/기구 /인원/시간/부대비용	재검률 1% ▼ = 1억 이상/月	

연구 배경

분야 소개

- 대량 생산체제: 스마트폰 1대당 카메라 모듈이 Dual, Triple, Quadruple로 생산량 증가
- 중요 관리사항: Particle 원소 추적
Particle 불량 검사

카메라 모듈의 주요 제조 공정

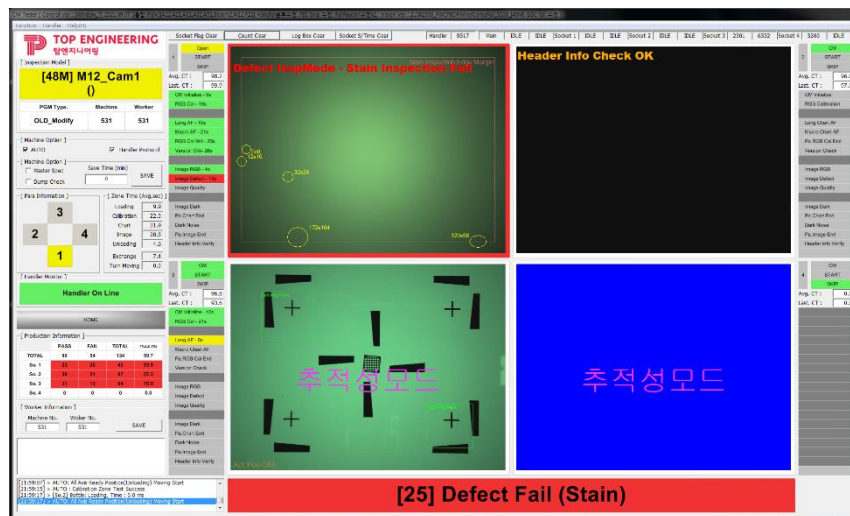


연구 배경

: 제품에 Particle 유입 → 흑점/얼룩 불량 발생


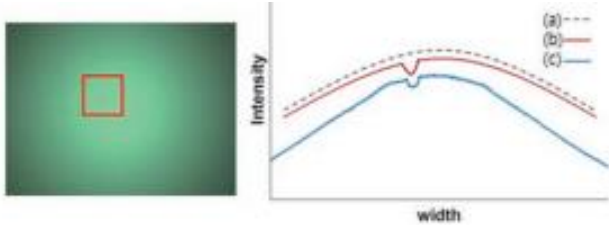
- 1) Backend 최종 검사장비는 전통적인 머신비전 기술로 이물불량을 판정 함
 - ① 정상품과의 비교
 - ② 픽셀 주변과의 밝기 차이
- 2) 오검으로 인한, 재검 비용 발생 초래

▶ 재검사를 최소화 하는, 새로운 검사방법 필요



기존 연구

■ 탐엔지니어링(자사 실제 장비)

항목	내 용
장비 제조사	(주)탐엔지니어링 - 검사항목: 제조검사 장비: Auto focus검사, Vignetting검사, 얼룩검사(1차)
연구 내용	<p>전자공학회지(2016.11)</p> <p>1. 머신비전 한계 : Lens shading(광량차이로 중심부는 밝고, 가장자리는 어두움)으로 인해 모두 같은 기준으로 검사할 수 없음</p> <p>※ 기존 비전검사 방식 ① 정상품과의 비교 ② 픽셀 주변과의 밝기 차이</p> <p>2. 연구 내용 - 머신러닝 기반 검출로 효과 기대</p> <div data-bbox="1537 551 1769 668">  <p>탐_기계학</p> </div> <div data-bbox="1078 831 1690 1053">  </div> <p>〈그림 6〉 Lens Shading (왼쪽) 카메라 모듈을 사용하여 촬영한 이미지. 박스 안에 원형 얼룩 존재. (오른쪽) 이미지의 X축 방향으로 밝기 그래프 (a)얼룩이 없을 경우 밝기 그래프 형태 (b)얼룩이 있을 경우 밝기 그래프 형태 (c)얼룩이 존재하는 지점의 실제 밝기 그래프</p>

딥러닝 모델을 통해, 불량 분류 정확도 비교

딥러닝 불량 판정(분류) 구현

- 데이터 수집: 4,800만 화소 카메라 모듈 검사 영상
- 영상 전처리: ①Resize, ②Data Augmentation(90/180/270 회전) = 24,836장
- 구현 모델/방법: RESNET18(Jupyter Notebook)

전처리: 2건

장비 검사 영상

①Resize

(1024*1024)

②Rotation

(90°/180°/270°)



(4,000*3,000)

CNN Network의 Input Data

원본		구분	Augmentation	
OK	NG		OK	NG
7,207	827	TRAIN(90%)	19,397	2,956
		TEST(10%)	2,155	328
8,034		계	21,552	3,284

Resnet18 Parameter

Data		Stochastic Gradient Descent		StepLR	
Set	Batch size	Learning Rate	Momentum	Step size	Gamma
24,836	4	0.001	0.9	7	0.1

연구 범위

딥러닝 분류 정확도 측정
(Resnet18)

정확도 비교

장비 판정정확도 측정
머신비전(양품비교, 주변밝기)

실험 결과

연구 결과: RESNET18(학습: 사전 학습모델 + 데이터셋)

Epoch 17/24

train Loss: 0.1613 Acc: 0.9572

test Loss: 0.5415 Acc: 0.8864

Epoch 18/24

train Loss: 0.1624 Acc: 0.9566

test Loss: 0.4355 Acc: 0.8913

Epoch 19/24

train Loss: 0.1618 Acc: 0.9570

test Loss: 0.4681 Acc: 0.8909

Epoch 20/24

train Loss: 0.1609 Acc: 0.9566

test Loss: 0.3873 Acc: 0.8933

Epoch 21/24

train Loss: 0.1622 Acc: 0.9571

test Loss: 0.4674 Acc: 0.8884

Epoch 22/24

train Loss: 0.1606 Acc: 0.9573

test Loss: 0.4608 Acc: 0.8925

Epoch 23/24

train Loss: 0.1616 Acc: 0.9571

test Loss: 0.4665 Acc: 0.8884

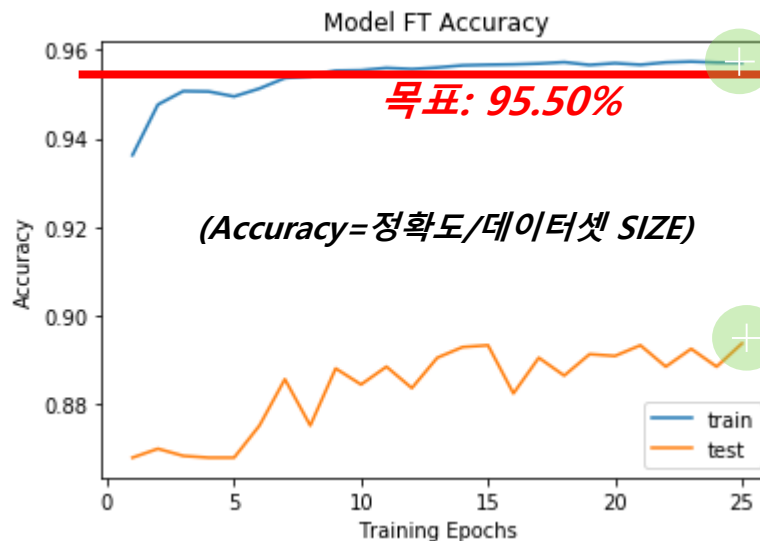
Epoch 24/24

train Loss: 0.1608 Acc: 0.9571

test Loss: 0.3895 Acc: 0.8937

학습 완료 걸린시간: 521분 14초
Best test Acc: 0.8937

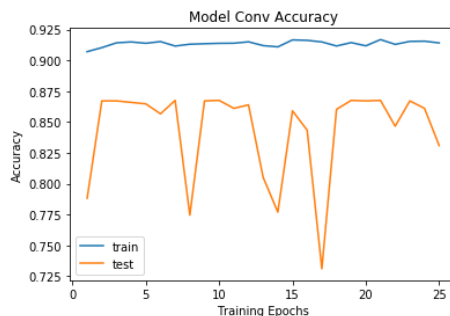
그림1. 실험 결과 그래프: 사전 학습모델 + 데이터셋



Train Best Acc.:
95.71%

Test Best Acc.:
89.37%

그림2. 별도 연구 결과: 데이터셋만 학습



■ 학습 결과

- 학습 완료 걸린시간: 216분 40초
- Best Train Acc: 91.68%
- Best Test Acc: 83.08%
- 학습시간은 낮으나 정확도 떨어짐

실험 결과

Confusion matrix: 실험 결과

Rule 기반 검사 - 머신 비전 장비

2021.11월 1개월 데이터

	양품률	불량률	
1차 판정	93.96%	6.04%	
최종 판정		양품 오검	진성 불량
	-	4.60%	1.44%

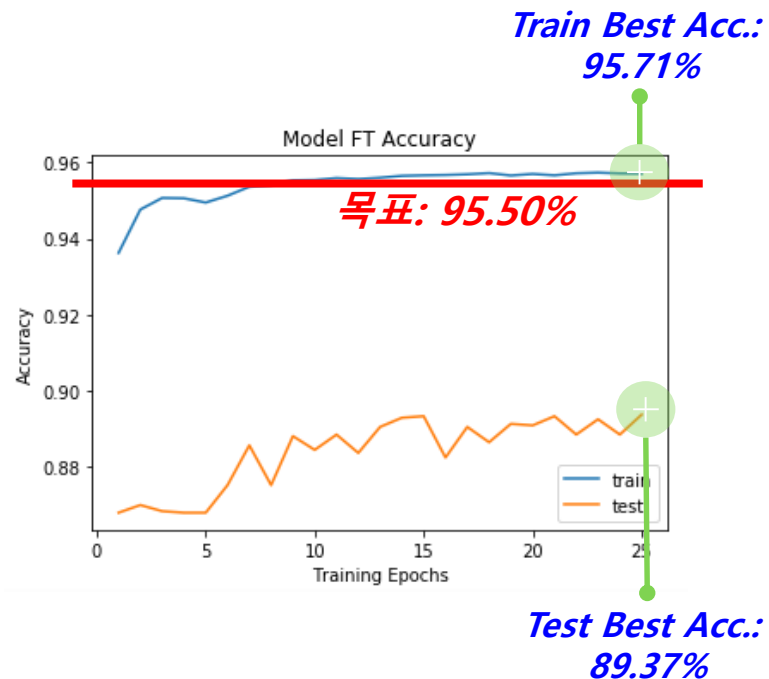
Confusion Matrix

		최종 판정 (%)	
		불량(T)	양품(F)
1차 판정 (%)	불량(T)	1.44	0.00
	양품(F)	4.60	93.96

구분	값
정확도(accuracy)	95.40%
재현율(recall)	100.00%
정밀도(precision)	23.84%
F1-Score	40.96%

딥러닝 기반 분류 - RESNET18

Accuracy = 정확도 / 데이터셋 SIZE)



연구의 의의 및 고찰

연구의 의의

- 룰(RULE) 기반 비전 검사 방식보다, 딥러닝 알고리즘 모델의 분류 정확도가 높음
- 학습 파라미터별 학습 시간과 정확도가 실제로 다양한 차이가 발생한다는 결과 도출
- 오(誤)검으로 인한 재검율을 낮출 수 있다면, 회사 이익에 상당한 기여

연구의 한계점 및 추가 연구 계획

- (한계) 데이터 불균형으로 인한, 정확도의 신뢰성 부족 불량 데이터 13%
- (한계) 영상이 고(高) 해상도로 인해, 데이터 셋의 크기와 학습시간이 높음
- (계획) 불량 데이터 셋 추가 확보하고, 학습 조건을 다양화하여 연구

결론 및 요약

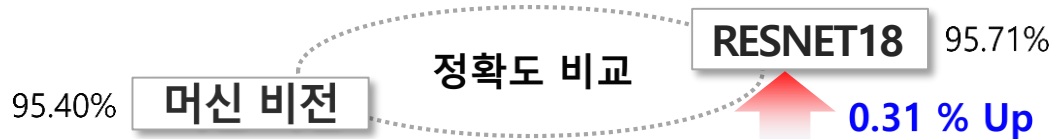
준비: 연구 준비에 따른, 결과 영향

- 데이터 셋: 양질의 데이터 수집이 상당히 중요하며, 이는 연구 결과와 직결 됨
- 알고리즘: 목적에 따른(Classification, Detection) 알고리즘 선정도 중요
- 학습 진행: 학습을 위한 전 처리와 파라미터에 따라 정확도 및 학습 효율에 영향

결과: RULE 기반 머신 비전 vs 딥러닝 불량 분류

- 딥러닝 모델 학습: RESNET18 사전 학습 모델 가중치와 데이터 셋을 함께 학습 진행
- 정확도
 - . 머신 비전: 1개월 간 생산 데이터를 Confusion Matrix로 정확도 기준을 잡음
 - . 딥 러닝: 분류 True 개수 / 전체 데이터 셋 개수

※ 결론: Train 데이터 기준 0.31 %가 높은 것으로 나타남



고찰

- 데이터 셋의 불균형(양품:불량=87%:13%)으로 인한, 정확도 신뢰도 제고 필요
- 다양한 알고리즘과 파라미터를 변화를 통한 실험 필요
- 학습시간을 줄이기 위한, 영상 Slicing 전처리 고민
- Detection까지 추가 연구 계획

감사합니다