

딥러닝과 영상처리를 활용한 무인계산시스템

김홍재, 최희웅, 윤보라, 김옥근, 조중휘
인천대학교 임베디드시스템공학과
khj40407, zika_zeno, qhfk1649, betrue0405 @naver.com
jcho@inu.ac.kr

1)

An Unmanned Checkout Counter using Deep Learning and Image Processing

Hongjae Kim, Heewoong Choi, Bora Youn, Okgeon Kim, Joongwhee Cho
Dept. of Embedded System Engineering, Incheon National University

요 약

본 논문은 대형마트와 같은 유통업계에서 무인계산시스템의 자동화로 소비자의 계산 편리성 증대를 위한 딥러닝과 영상처리를 활용한 무인계산시스템을 제안한다. 소비자가 무인계산시스템의 컨베이어 벨트 위에 계산할 물품을 올리면 벨트 끝에 위치한 카메라로 이동하여 촬영한 물품의 이미지를 딥러닝과 영상처리로 분석하여 제품의 리스트를 제공, 결제가 완료되면 서버에 전송하여 재고를 관리하고 발주가 필요한 제품은 자동으로 발주하는 시스템이다.

1. 서론

최저임금의 급격한 인상과 경기침체로 유통업체들의 비용 부담이 늘어나면서 종업원을 줄이는 ‘무인 자동화 시스템’이 본격화되고 있다.

본 연구는 위와 같은 배경에서 사용자에게는 더 나은 서비스를 제공하고 기업에게는 더 효율적이고 비용을 절약하게 하기 위해서 이다. 계산을 담당하는 계산원의 수는 줄어들고 전체 시스템을 관리하는 소수의 관리자를 두어 노동인구 감소에 따른 인력난에 능동적으로 대체할 수 있게 해주는 무인계산시스템을 제안한다. 무인계산시스템은 영상처리, 딥러닝 기술을 활용하여 마트 이용객이 어떤 상품을 얼마큼 컨베이어에 담았는지 측정하고 각 상품에 대한 가격을 모니터에 출력하여 사용자가 빠른 결제를 할 수 있도록 유도한다.

2. 기술적 내용

2.1 제안하는 시스템 설계

딥러닝을 활용한 무인 계산 시스템의 전체적인 흐름은 그림 1과 같이 구성한다.

*이 논문은 2018 한이음 ICT 멘토링 프로젝트의 연구비 지원을 받은 결과물입니다.



(그림1) 전체적인 흐름

딥러닝을 활용한 무인 계산 시스템을 구성하는 핵심 컴포넌트는 다음과 같은 단계를 거친다.

2.2 상품 이미지 학습과 추출 및 분류

본 연구에서는 CNN 모델 중 'Alexnet'을 이용하였고, Alexnet의 네트워크 아키텍처는 그림 2와 같이 구성다.

네트워크에는 5 개의 컨벌루션 레이어와 3 개의 완전히 연결된 레이어가 있다. 16 레이어까지 추출을 하며 17 레이어부터 25 레이어까지 검출을 한다. 상품 이미지가 다섯 번의 Convolution Layer 와 두 번의 Dropout Layer 를 거치며 특징 추출(Feature Extraction)을 진행하고, 마지막엔 Softmax Layer 를 통과하여 분류(Classification)가 된다.

```

Layers =
25x1 Layer array with layers:
 1 'data' Image Input 227x227x3 images with 'zerocenter' normalization
 2 'conv1' Convolution 96 11x11x3 convolutions with stride [4 4] and padding [0 0 0 0]
 3 'relu1' ReLU
 4 'norm1' Cross Channel Normalization cross channel normalization with 5 channels per element
 5 'pool1' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
 6 'conv2' Convolution 256 5x5x48 convolutions with stride [1 1] and padding [2 2 2 2]
 7 'relu2' ReLU
 8 'norm2' Cross Channel Normalization cross channel normalization with 5 channels per element
 9 'pool2' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
10 'conv3' Convolution 384 3x3x256 convolutions with stride [1 1] and padding [1 1 1 1]
11 'relu3' ReLU
12 'conv4' Convolution 384 3x3x192 convolutions with stride [1 1] and padding [1 1 1 1]
13 'relu4' ReLU
14 'conv5' Convolution 256 3x3x192 convolutions with stride [1 1] and padding [1 1 1 1]
15 'relu5' ReLU
16 'pool5' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
17 'fc6' Fully Connected 4096 fully connected layer
18 'relu6' ReLU
19 'drop6' Dropout 50% dropout
20 'fc7' Fully Connected 4096 fully connected layer
21 'relu7' ReLU
22 'drop7' Dropout 50% dropout
23 'fc8' Fully Connected 1000 fully connected layer
24 'prob' Softmax softmax
25 'output' Classification Output crossentropyx with 'tench', 'goldfish', and 998 other classes
    
```

(그림 2) alexnet network with architecture

3. 구현과 실험을 통한 인공신경망(ANN) 성능 평가

실제 유통, 판매 되고있는 상품 5개로 신경망 트레이닝을 진행하였다. 각 상품당 이미지 데이터는 1200~1500장이며, 학습단위인 미니 배치 사이즈(mini-batch size)는 64이므로 반복되는 학습단위로 가져오는 이미지는 64장이다.

그림 3은 설정 값에 따른 인공 신경망의 성능을 나타 낸 그림이다. 여기서 에폭을 20을 설정하였고, 트레이닝 결과 유효성 98%의 평균 정확도를 보였다.

Epoch	Iteration	Time Elapsed (seconds)	Mini-batch Loss	Mini-batch Accuracy	Base Learning Rate
1	1	4.07	2.3903	14.06%	0.0010
1	50	30.94	0.0001	100.00%	0.0010
2	100	58.26	0.0000	100.00%	0.0010
2	150	85.57	0.0004	100.00%	0.0010
3	200	112.89	0.0000	100.00%	0.0010
3	250	140.09	0.0001	100.00%	0.0010
4	300	167.34	0.0002	100.00%	0.0010
4	350	194.64	0.0001	100.00%	0.0010
5	400	221.94	0.0001	100.00%	0.0010
5	450	249.17	0.0000	100.00%	0.0010
6	500	276.43	0.0001	100.00%	0.0010
6	550	303.80	0.0000	100.00%	0.0010
7	600	331.05	0.0000	100.00%	0.0010
8	650	358.30	0.0001	100.00%	0.0010
8	700	385.53	0.0000	100.00%	0.0010
9	750	412.77	0.0000	100.00%	0.0010
9	800	440.06	0.0000	100.00%	0.0010
10	850	467.52	0.0001	100.00%	0.0010
10	900	494.76	0.0000	100.00%	0.0010
11	950	522.08	0.0000	100.00%	0.0010
11	1000	549.36	0.0000	100.00%	0.0010
12	1050	576.68	0.0000	100.00%	0.0010
12	1100	603.95	0.0000	100.00%	0.0010
13	1150	631.39	0.0000	100.00%	0.0010
14	1200	658.85	0.0000	100.00%	0.0010
14	1250	686.13	0.0000	100.00%	0.0010
15	1300	713.48	0.0000	100.00%	0.0010
15	1350	740.90	0.0002	100.00%	0.0010
16	1400	768.15	0.0000	100.00%	0.0010
16	1450	795.42	0.0000	100.00%	0.0010
17	1500	822.68	0.0000	100.00%	0.0010
17	1550	849.99	0.0000	100.00%	0.0010
18	1600	877.68	0.0000	100.00%	0.0010
18	1650	905.22	0.0000	100.00%	0.0010
19	1700	934.19	0.0000	100.00%	0.0010
20	1750	962.12	0.0000	100.00%	0.0010
20	1800	989.62	0.0000	100.00%	0.0010
20	1840	1011.53	0.0001	100.00%	0.0010

(그림 3) epoch

4. 가치 및 활용방안

본 무인계산시스템은 전체적으로 정확한 상품의 판별과 빠른 처리 속도를 보여주었다. 이는 현재 직원이 계산하는 것보다 빨랐다는 것을 의미하며, 실제로 각 매장에 시스템이 적용이 된다면 직원이 줄어들어 가게가 부담해야할 노동력은 줄어들어 매출이 증대될 것이다. 또한 시스템이 보급된다면 이 이 시스템을 양산하기 위한 기업의 필요 노동력은 증가하게 되므로 부족한 일자리 창출에도 기여 할 수 있게 된다. 인력난이 예상되는 지역을 중심으로 보급된다거나 반대로 사람이 많이 사는 주요 도심지(계산대에 대기자가 많아 대기 시간이 길어질 것으로 예상되는 곳에서 더 높은 수준의 서비스를 이용자에게 공급하게 될 것이다.

5. 결론 및 향후연구

실 매장에 적용된다면 현재 무인결제시스템 도입 초기이기 때문에 기기 사용 미숙으로 불편함을 느낄 수 있지만 제안한 무인결제시스템은 사용자에게 작동을 요하는 작업이 현저히 줄어들어 기기 사용에 편리성이 향상될 것으로 예상된다[1]. 현재 무인계산대의 보급율이 증가함에 따라 딥러닝과 영상처리를 활용한 무인계산시스템의 향후 수요가 증가할 것과 소비자의 만족도 또한 증가할 것으로 기대된다[6].

개발한 시스템을 구현해 제품의 이미지 분석 결과 98%의 정확도와 4초/개의 처리 속도를 얻었다. 다만 다른 제품의 각도에 따른 같은 이미지인 경우와 1차 상품(농림수산물 등의 가공도가 높아지기 이전의 원료에 가까운 단계의 생산물) 등 이미지 미분류, 미인식 상품을 보완할 기술이 필요하여 지속적인 연구와 개발로 미분류, 미인식 상품의 보완과 분류 정확성과 처리 속도를 향상시켜 실생활에서 사용자에게 유용한 시스템을 제공할 계획이다.

참고문헌

- [1]권용철, 김관배 (2015,)구성요소의 특성에 따른 무인계산대 사용성 연구, 한국디지털디자인협의회, 디지털디자인학연구 15(2), 637-648
- [2]김민성, 김영민, 박태수 (2013), 최저임금 변화가 지역고용에 미치는 효과분석, 한국고용노사관계학회, 산업관계연구 23(2), 37-73
- [3]고성제, 김영현 (2017), 딥러닝을 이용한 객체분류 및 검출 기술, 대한전자공학회, 전자공학회지 44(11), 26-33
- [4]조재현, 장영민, 조상복(2015), 다중 ROI에서 영상 화질 표준화 및 선택적 허프 변환 알고리즘을 통한 고성능의 차선 인식 알고리즘, 대한전자공학회, 전자공학회논문지 52(2), 148-161
- [5]이대회, 이재용, 배규현, 김기형, 노중호 (2018), 스마트팩토리를 위한 실시간 자동발주 Tool매니지먼트 시스템, 한국정밀공학회, 한국정밀공학회 학술발표대회 논문집, 418-419
- [6]강진희 (2018), 외식기업의 무인주문결제시스템에 대한 소비자 수용의도 연구 : 챗봇과 디지털 키오스크를 중심으로, 한국관광연구학회, 관광연구저널 32(1), 153-168