

정확한 돼지 탐지를 위한 박스 레벨 후처리

유승현*, 김대원*, 김희곤*, 정용화*, 박대희*

*고려대학교 컴퓨터융합소프트웨어학과

e-mail: tidslsld44@korea.ac.kr

Box-Level Post Processing for Accurate Pig Detection

SeungHyun Yu*, Daewon Kim*, Heegon Kim*, Yongwha Chung*, and
Daihee Park*

*Dept. of Computer Convergence Software, Korea University

요 약

합성곱 신경망(Convolutional Neural Network) 기반 객체 탐지기의 발전으로 돈사에서 돼지 모니터링이 가능하지만, 실제 농가에서 적용하기 위해서는 탐지기의 정확도를 개선해야 하는 문제가 여전히 남아있다. 본 연구에서는 합성곱 신경망 기반 돼지 탐지기의 출력인 박스들의 신뢰도 값을 평가하고 잘못된 박스들의 신뢰도 값을 보정하는 박스 레벨 후처리 방법을 제안한다. 즉, 신뢰도 값이 가짜 돼지인지 진짜 돼지인지 애매한 경우, 박스내 전경 픽셀 정보와 인접 박스의 정보를 이용하여 신뢰도 값을 낮추거나 높이는 보정 작업을 수행한다. 그리고 실제 돈사에서 취득한 11,308장의 영상 데이터로 실험한 결과, 제안 방법은 합성곱 신경망 기반 돼지 탐지기의 에러율을 4.4%에서 1.2%로 개선하는 효과가 있음을 확인하였다.

1. 서론

돈사내 작업자의 부족(국내의 경우 작업자 1명이 평균 2,000 마리의 돼지를 관리)과 돼지의 높은 폐사율(국내의 경우 연간 약 500만 마리의 돼지가 폐사)을 고려할 때, 개별 돼지에 대한 세밀한 관리를 위하여 정보기술(Information Technology)을 적용한 돈사 모니터링의 필요성이 증가하고 있다[1]. 그러나 지속적으로 발전하는 합성곱 신경망 기반 객체 탐지 기술[2]을 적용하여도 돼지들 간의 겹침(occlusion) 등의 이유로 혼잡한 돈사 내 돼지들을 정확히 탐지하는데 한계가 있다. 본 연구에서는 이러한 어려움을 극복하기 위하여 합성곱 신경망 기반 돼지 탐지기의 출력인 박스들의 신뢰도 값을 평가하고 잘못된 박스들의 신뢰도 값을 보정하는 박스 레벨 후처리 방법을 제안한다.

본 연구를 위하여, 돈사 내 위에서부터 아래로 보는(top-view) 카메라로부터 획득된 영상을 최근 발표된 YOLOv4[3] 객체 탐지기에 적용하여 탐지된 객체에 대응하는 바운딩 박스(bounding box)들을 획득한다. YOLOv4는 2020년 4월 발표일 기준으로, 대표적인 객체 탐지 공개 DB인 MS COCO[4]로 처리속도 대비 가장 높은 정확도를 달성하여 본 연구의 객체 탐지기로 결정하였다. 즉, 합성곱 신경망 기반 딥러닝 기술의 지속적인 발전으로 YOLOv4는 대부분의 돼지들을 정확히 탐지하지만, 돈방의 복잡한 구조 등의 이유로 배경(background)에 해당하는 가짜 돼지를 탐지한 박스들이 포함되는 문제(False Positive, FP)가 남아있다. 만약 이러한 가짜 돼지들을 제거하기 위하여 단순히 박스의 신뢰도(confidence) 값이 적

은 것을 제거한다면, 겹친 돼지(occluded pig)와 같이 신뢰도 값이 낮은 다른 박스도 함께 제거된다는 또 다른 문제(False Negative, FN)가 발생한다. 이러한 문제들을 해결하기 위하여 본 연구에서는 현재 박스의 신뢰도 값으로는 가짜 돼지인지 진짜 돼지인지 판단하기 애매한 경우, 박스내 전경(foreground) 픽셀 정보와 인접 박스의 정보를 이용하여 신뢰도 값을 보정하는 작업을 수행한다. 그리고 실제 돈사에서 취득한 영상 데이터로 제안 방법의 효과를 확인한다.

2. 제안 방법

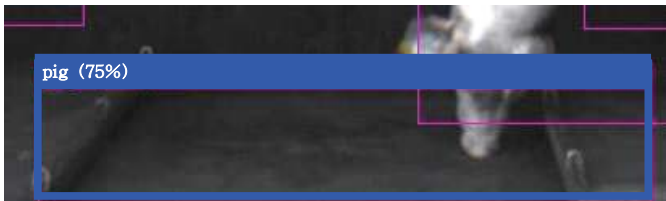
본 연구에서는 9마리의 돼지가 있는 돈방에 위에서부터 아래로 보는 카메라를 설치하여 영상을 획득하고, 이를 최근 발표된 YOLOv4 객체 탐지기에 적용하여 박스들을 획득한다고 가정한다. 합성곱 신경망 기반 딥러닝 기술의 지속적인 발전으로 YOLOv4는 대부분의 돼지들을 정확히 탐지(박스의 신뢰도 값이 80% 이상)하지만, 돈방의 복잡한 구조와 혼잡한 돈방 내 돼지들의 겹침으로 가짜 돼지 박스가 진짜 돼지 박스 보다 높은 신뢰도 값을 갖는 경우가 발생한다. 본 연구에서는 이러한 문제를 해결하기 위하여, 박스내 전경 픽셀 정보를 이용하여 가짜 돼지의 신뢰도 값을 낮추고 인접 박스의 정보를 이용하여 겹친 돼지의 신뢰도 값을 높이는 보정 방법을 제안한다.

2.1 전경 픽셀 정보를 이용한 신뢰도 보정(Decrease_Confidence)

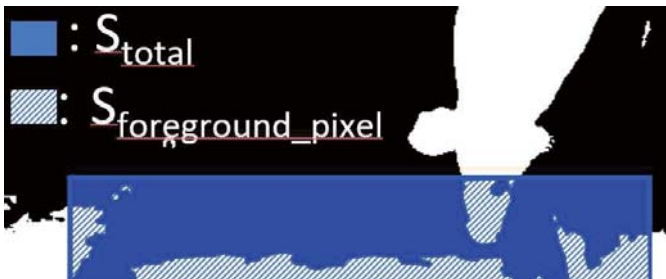
먼저 신뢰도 값이 애매한 범위에 있는 박스가 배경에

해당하는 가짜 돼지 박스인지를 판단하기 위하여, 해당 박스에 전경 픽셀이 얼마나 포함되는지 계산한다. 돼지의 위치에 따라 박스 내 전경 픽셀 비율이 다르긴 하지만, 가짜 돼지에 대응하는 박스라면 절반 이하의 픽셀이 전경 픽셀이라 가정한다. 박스 내 픽셀들을 전경과 배경으로 구분하기 위하여 적응 임계 값(adaptive threshold) 방법[5]을 적용하고, 이러한 기준으로 판단된 박스의 신뢰도 값을 해당 박스의 전체 크기(S_{total}) 대비 해당 박스의 전경 영역의 크기($S_{foreground_pixel}$)를 이용하여 다음의 수식으로 보정한다.

$$dified_Confidenece_Score = Confidence_Score - (100 - Confidence_Score) * \frac{S_{foreground_pixel}}{S_{total}} \quad (1)$$



(그림 1) YOLOv4의 가짜 돼지 탐지 예



(그림 2) 전경 픽셀 정보를 이용하여 전경 영역의 크기($S_{foreground_pixel}$) 계산



(그림 3) 제안 방법을 적용한 가짜 돼지 보정 결과

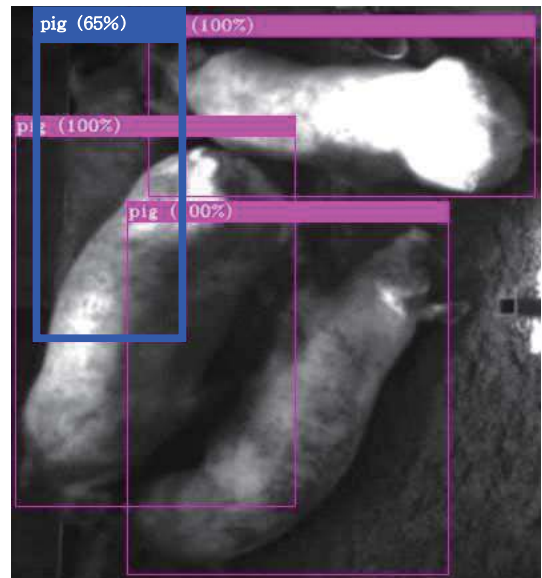
그림 1에 이러한 경우의 예를 보여주는데, 그림 2와 같이 전경 픽셀 정보를 이용하여 가짜 돼지 박스의 신뢰도 값이 최초 75%에서 63%로 보정됨을 확인(그림 3)할 수 있다.

2.2 인접 박스 정보를 이용한 신뢰도 보정

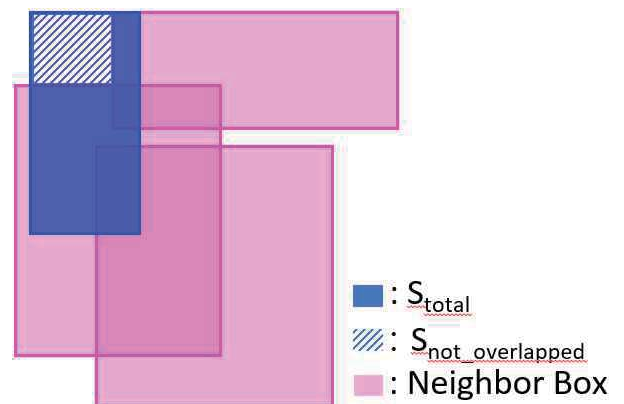
(Increase_Confidence)

다음으로 신뢰도 값이 애매한 범위에 있는 박스가 전경에 해당하는 겹친 돼지 박스인지를 판단하기 위하여, 해당 박스와 겹치는 박스가 존재하는지 판단한다. 참고로 이러한 경우라면 해당 박스 내 전경 픽셀 비율이 절반 이상이 될 것이고, 앞선 2.1에서 언급한 신뢰도 값을 낮추는 보정 과정이 적용되지 않을 것이다. 즉, 겹친 돼지의 경우 인접 돼지에 의하여 신뢰도 값이 낮게 계산되었다고 가정할 수 있으며, 이러한 가정을 기반으로 해당 박스의 전체 크기(S_{total}) 대비 해당 박스의 겹치지 않은 영역의 크기($S_{not_overlapped}$)를 이용하여 다음의 수식으로 보정한다.

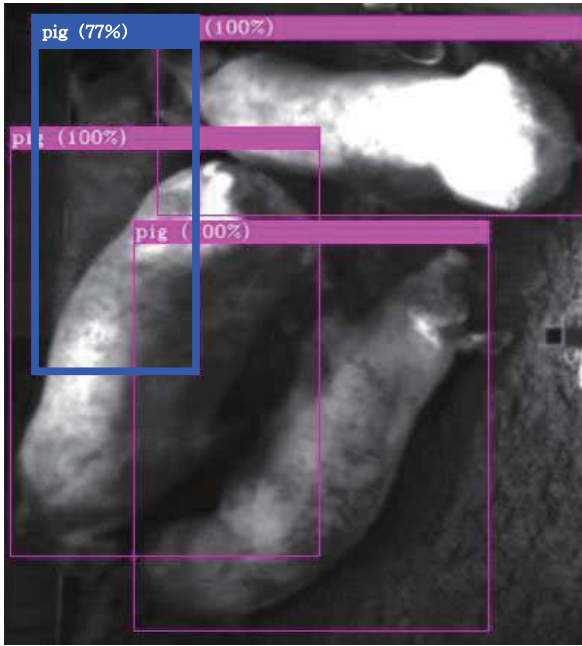
$$Modified_Confidenece_Score = Confidence_Score + (100 - Confidence_Score) * \frac{S_{not_overlapped}}{S_{total}} \quad (2)$$



(그림 4) YOLOv4의 겹친 돼지 탐지 예



(그림 5) 인접 박스 정보를 이용하여 겹치지 않은 영역의 크기($S_{not_overlapped}$) 계산



(그림 6) 제안 방법을 적용한 겹친 돼지 보정 결과

그림 4에 이러한 경우의 예를 보여주는데, 그림 5와 같이 인접 박스 정보를 이용하여 겹친 돼지 박스의 신뢰도 값이 최소 65%에서 77%로 보정됨을 확인(그림 6)할 수 있다. 또한, 전경 픽셀 정보와 인접 박스 정보를 이용한 신뢰도 값의 전체 보정 과정을 요약하면 **알고리즘 1**과 같다.

Algorithm 1. 박스의 신뢰도 값 조절

입력

YOLOv4에서부터 박스들의 집합인 B 를 받는다.
박스 신뢰도의 기준점을 T 라 칭한다.

출력

최종 탐지에 사용할 B 를 알고리즘 1을 통해서 구한다.

Algorithm

$B' = B$;

B 를 신뢰도에 따라서 오름차순으로 분류한다.

for each box b in B' do

if confidence[b] < $T-10$ then

b 를 B' 로부터 지운다.

else if $T \leq \text{confidence}[b] < T+10$ then

Decrease_Confidence(b) 함수를 호출

else if $T-10 \leq \text{confidence}[b] < T$ then

Increase_Confidence(b) 함수를 호출

if confidence[b] < T then

b 를 B' 에서 지운다.

Return B' ;

3. 실험 결과

본 실험은 충북대학교내 9마리 돼지(110~115일령, 90~98kg, Duroc × Landrace × Yorkshire)가 있는 돈사에서 획득된 영상으로 수행되었다. 영상 촬영을 위해 돈사의

중앙을 기준으로 3.2m 높이의 천장에 Intel® RealSense D435 카메라를 설치하였고, 이를 통해 (1280 × 720) 해상도의 돼지 영상을 초당 30장의 속도로 획득하였다. 또한 첫날 수집된 영상으로부터 임의의 582장을 선별하여 YOLOv4를 학습(영상 해상도: (416 × 416), 학습 횟수: 10,000회, 학습률: 0.0001)하였고, 다음날 수집된 2시간 분량의 영상(총 216,000장)에서 프레임 차를 적용하여 움직임에 대표하는 11,308장의 키 프레임에 선별한 후 테스트하였다. 마지막으로, 본 실험은 Intel® Core i7-7700 3.60 GHz CPU, 12GB RAM, Visual Studio 2017, OpenCV 3.4, GeForce RTX 2080Ti GPU, CUDA 10.0 환경에서 수행되었다.

본 실험에서는 돈사 내 9마리의 돼지가 있다는 사실을 가정하여, YOLOv4가 10개 이상의 박스를 출력하는 경우에는 신뢰도 값을 기준으로 상위 9개의 박스만 선택하는 추가 과정을 적용하여 YOLOv4의 최종 결과로 활용하였다. 반면, 돼지들이 완전히 겹쳐 육안으로는 영상 내 겹친 돼지들의 마릿수를 판단할 수 없는 경우도 있어, 9개보다 적은 수의 박스가 탐지되는 경우에 상위 9개의 박스를 선택하는 추가 과정을 적용하지는 않았다. 또한, 제안 방법에서는 최종적으로 신뢰도 값이 70% 이상인 박스들만 최종 결과로 활용(즉, $T = 70\%$)하였다. 이는 신뢰도 값을 40~90%로 하여 각각 실험한 결과, $T=70\%$ 로 설정하는 것이 가장 좋다고 판단이 되어 $T=70\%$ 로 설정하였다. 또한, 정확도(ACC)는 다음과 같이 계산하였다.

$$CC = \frac{TP}{TP + FN + FP} \quad (3)$$

성능 측정 결과, 제안 방법은 YOLOv4의 95.6% 정확도를 98.8%로 개선할 수 있음을 확인하였다. 또한, 제안 방법의 후처리 과정에 소요되는 오버헤드 시간이 무시할 정도이며, YOLOv4를 포함한 전체 수행 과정이 실시간에 처리될 수 있음을 확인하였다(표 1 참조).

<표 1> 기존 방법과 제안 방법의 성능 비교

	정확도(%)	수행시간(msec)
YOLOv4[3]	95.6	11.21
YOLOv4 +제안 방법	98.8	11.36

4. 결론

본 연구에서는 YOLOv4의 출력인 박스들의 신뢰도 값을 평가하고 보정하여, 돼지 탐지 정확도를 개선하는 박스 레벨 후처리 방법을 제안하였다. 즉, 대부분의 박스가 진짜 돼지(전경)인 경우는 신뢰도 값이 높고 가짜 돼지(배경)인 경우 신뢰도 값이 낮아 문제가 없지만, 애매한 신뢰도 값을 갖는 박스에 대하여 보다 세밀한 평가가 필요하였다. 따라서 이러한 박스들에 대하여, 박스내 전경 픽셀 정보와 인접 박스의 정보를 이용하여 신뢰도 값을 낮추거

나 높이는 보정 작업을 수행하였다.

제안 방법의 효과를 확인하기 위해 돈방 내 9마리 돼지가 있는 실제 돈사에서 취득한 11,308장의 영상 데이터로 실험한 결과, 제안 방법은 YOLOv4 기반 돼지 탐지기의 오답률을 4.4%에서 1.2%로 개선하는 효과가 있음을 확인하였다. 추가로 돈방의 비디오 모니터링 특성을 활용한다면 폐쇄된 돈방(closed pig room)내 돼지 탐지 정확도를 보다 개선할 수 있을 것으로 기대된다.

감사의 글

이 논문은 2018년도 교육부의 재원으로 한국연구재단의 기초연구사업(2018R1D1A1A09081924)과 2019년도 과학기술정보통신부의 재원으로 한국연구재단-현장맞춤형 이공계 인재양성 지원사업(No. 2019H1D8A1109907)의 지원을 받아 수행된 연구임.

참고 문헌

- [1] S. Matthews, et al., "Early Detection of Health and Welfare Compromises through Automated Detection of Behavioural Changes in Pigs," *The Veterinary Journal*, Vol. 217, pp. 43-51, 2016.
- [2] L. Liu, et al., "Deep Learning for Generic Object Detection: A Survey," *International Journal of Computer Vision*, Vol. 128, pp. 261-318, 2020.
- [3] A. Bochkovskiy, C. Wang, and H. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.
- [4] T. Lin, et al., "Microsoft COCO: Common Objects in Context," *In Proceedings of the European Conference on Computer Vision(ECCV)*, 2014.
- [5] I. Blayvas, A. Bruckstein, and R. Kimmel, "Efficient Computation of Adaptive Threshold Surfaces for Image Binarization," *Pattern Recognition*, Vol. 18, No. 1, pp. 89-101, 2006.