

영상기반 축산동물 관리를 위한 회전형 동물검출 신경망 연구

전 광 명[°], 안 형 준^{*}, 주 소 현^{*}, 류 인 철^{*}, 권 슬 빈^{*}, 정 진 우^{*}, 이 용 기^{**}

A Study on Rotated Animal Detection Neural Network for Image-Based Livestock Management

Kwang Myung Jeon[°], HyungJun Ahn^{*}, Soheun Ju^{*}, Inchul Ryu^{*},
 SolBeen Kwon^{*}, Jinwoo Jung^{*}, Younggie Lee^{**}

요 약

본 논문에서는 영상기반의 축산동물 관리를 위한 회전형 동물검출신경망을 제안한다. 제안하는 신경망은 영상 내 축산동물 각각의 검출 영역을 정확하게 정의하기 위해 동물의 몸통의 폭과 길이, 그리고 회전 방향을 추정한다. 단순히 검출 영역을 직사각형을 표현하는 기존방법과 달리 제안하는 신경망은 해당 동물이 영상 내 어느 방향으로 서 있어도 해당 동물의 몸통을 제외한 주변 영역의 지정을 최소화하여 다수의 동물이 밀집한 경우와 같은 보편적 축산동물 관리에 사용하기 적합하다. 제안하는 방법의 성능을 검증하기 위해, 7개의 농장과 1개의 가상생성 농장에서 취득한 돈사 내부의 이미지 26,208장을 준비하였다. 이 데이터를 활용하여 기존의 직사각형 출력 형식을 지니는 방법과 제안방법을 학습한 후, 학습에 활용되지 않은 이미지 1,333장을 이용하여 평가를 수행하였다. 평가결과 제안방법은 0.9534의 높은 mAP50 수치를 보이며, 기존방법 대비 9.04% 높은 동물검출 성능을 보였다. 또한, 제안하는 신경망은 13.0GFLOPS의 연산량을 지니며 128-core Maxwell GPU를 지니는 임베디드 장치에서 29.87 FPS의 처리속도로 구현될 수 있음을 확인하였다. 이는 제안하는 신경망이 인터넷이 끊긴 축산현장에서도 엣지컴퓨팅 형태로 구현 가능함을 의미한다.

Key Words : Object Detection, Artificial Intelligence, SmartFarm, Livestock Management, Computer Vision

ABSTRACT

This paper introduces a rotated animal detection neural network for livestock management utilizing image analysis. The neural network estimates the width, length, and rotation direction of the animal's torso to accurately define the detection area of each individual livestock animal in the image. Unlike the existing method that expresses the detection area as a rectangle, our approach minimizes the surrounding area except for the animal's body, regardless of its direction in the image. This makes it applicable in livestock management environments with densely packed animals. We prepared 26,208 images of pig farms from 7 farms and 1 virtual farm to verify our proposed method. We trained our proposed method and the existing

* 본 논문은 농림축산식품부 및 과학기술정보통신부, 농촌진흥청의 재원으로 농림식품기술기획평가원과 재단법인 스마트팜연구개발사업단의 스마트팜다부처패키지혁신기술개발사업의 지원을 받아 연구되었음 (421015-04)

[°] First and Corresponding Author : AI Convergence Technology Laboratory, Intflow Inc., kmjeon@intflow.ai, 정희원

^{*} AI Convergence Technology Laboratory, Intflow Inc.

^{**} Dept. of Animal Resource, Chonnam National University

논문번호 : 202307-146-0-SE, Received June 1, 2023; Revised August 23, 2023; Accepted August 23, 2023

rectangular output format using this data and evaluated both using 1,333 unseen/untrained images. Our proposed method showed an mAP50 value of 0.9534, outperforming 9.04% better than the conventional method. Additionally, we confirmed that the proposed neural network has a low computational load of 13.0 GFLOPS and can be implemented at a processing rate of 29.87 FPS on an embedded device with a Maxwell GPU with 128 cores. Therefore, our proposed neural network can be effectively employed in an edge computing paradigm, even at livestock sites without internet access.

I. 서 론

축산업은 인류의 식량공급의 33%를 차지하는 중요한 1차산업 중 하나이다^[1]. 대표적인 축산업 분야는 육우, 낙농, 양돈, 양계로 나눌 수 있다. 각 축종별 축산업의 공통적인 특성은 소수의 작업자가 절대다수의 동물을 관리한다는 점이다^[1]. 다수의 동물을 효과적으로 관리하기 위해 축종별로 다양한 정보통신 기술 (ICT: Information Communication Technology) 이 제안되었다^[2-4]. 그간의 축산동물 관리 기술은 주로 동물의 몸에 센서를 부착시켜 정보를 취득하는 접촉식 센싱 장치 형태로 발전되어 왔다^[2-4]. 대표적으로 상업화된 방법 중 하나로 축우 개체관리를 위한 목걸이형 관리장치를 들 수 있다^[2,3]. 목걸이형 관리장치는 가속도계센서를 활용하여 활동량이나 이상행동 등을 인지하여 개체별 건강 상태를 측정할 수 있으며, 특히 질병 조기진단 및 발정 탐지 등에 유용하게 활용되어 농장의 생산성 개선에 도움을 주는 것으로 알려져 있다^[2,3]. 하지만 이러한 접촉식 센싱 장치는 축산업 다방면에 보편적으로 확산되지 못한 상황인데, 다음과 같은 한계가 존재하기 때문이다^[4]. 1) 접촉식 센서에 내장된 배터리의 잠재적 위해성이 존재한다. 2) 배터리수명의 한계에 따른 효율성이 저하된다. 3) 축종당 단가 대비 접촉식 센싱 장치의 제조비용이 상대적으로 높은 수준이다. 4) 축산동물이 출하순환에 따라 장치의 탈착 및 재부착이 번거롭다. 5) 동물의 성향으로 인한 접촉식 센싱 장치의 파손이 발생할 수 있다. 따라서 접촉식 센싱 장치의 단점들을 극복하기 위한 새로운 축산동물 관리 기술이 요구되는 실정이다. 최근 3년간 접촉식 센싱 방식의 한계를 극복하기 위한 비디오 카메라를 사용한 영상기반의 비접촉 동물관리 기술들이 활발히 연구되고 있다^[5,6]. 동물 행동 및 복지 모니터링을 위한 비디오 카메라의 사용은 동물 건강 문제의 더 나은 감지, 사육동물 이동의 개선 및 동물과 작업자의 안전성 향상과 같은 중요한 이점을 제공할 수 있다^[6]. 그러나 영상기반 시스템을 사용한 대규모 사육 업무를 효과적으로 관리하려면 동물 감지 및 추적에 대한 정교한 알고리즘이 필요하다.

동물 감지 및 추적을 위한 기존의 영상기반 사육 관리 시스템 방법은 주로 사각형 경계 상자 (rectangular bounding box)를 사용하여 이미지에서 동물의 위치를 정의하는 보편적 객체 감지 알고리즘을 기반으로 한다^[7,8]. 그러나 이러한 방법에는 여러 제약 사항이 있다. 첫째, 사각형 경계 상자는 영상 내 동물의 몸통이 화각 내 축과 정렬되지 않은 경우, 실제 몸통에 해당하지 않는 배경 영역까지 과대 추정하여 검출 및 추적 오류를 발생시킬 수 있다. 둘째, 동물이 밀집되어 있거나 카메라 각도가 변경될 때 사각형 경계 상자는 높은 수준의 거짓 양성 (false-positive) 과 거짓 음성 (false-negative) 을 유발할 수 있다. 마지막으로, 축산동물 검출 정확성을 높이기 위한 키포인트^[9]와 폴리곤^[10] 등 추가 정보를 활용하는 모델이 있지만 추론 복잡도가 증가하는 단점이 있다.

이러한 한계를 극복하기 위해 본 논문에서는 영상기반 사육 관리를 위한 회전형 동물검출신경망을 제안한다. 제안하는 신경망은 동물의 몸통 너비, 길이 및 회전 방향을 추정하여 각 동물의 감지 영역을 정확하게 정의한다. 기존 방법은 감지 영역을 단순히 사각형으로 표시하지만, 제안된 신경망은 동물이 동영상에서 어느 방향으로 서 있든 몸통을 제외한 주변 영역을 최소화한다. 따라서 다수의 동물이 밀집한 때도 검출 영역을 정확하게 표현하고 모호성을 최소화한다.

제안된 신경망은 기존 방법보다 여러 가지 이점이 있다. 첫째, 영상 속 다양한 방향과 위치에 있는 동물의 영역을 회전 경계상자로 정확히 검출하고 추적할 수 있어, 이로 인해 false positives와 false negatives가 유의미하게 감소된다. 둘째, 밀집한 환경에서 다수의 동물을 감지하고 추적할 수 있어 대규모 축산동물 관리에 적합하다. 마지막으로 키포인트나 폴리곤 등 복잡한 메타정보 없이 구현되어 사각형 경계 상자 방식 대비 추론 복잡도가 크게 증가하지 않는다.

이 논문의 나머지 부분은 다음과 같이 구성된다. 2장에서는 제안된 회전형 동물검출 신경망을 자세히 설명한다. 3장에서는 제안된 방법의 실험 결과를 제시하고 기존 방법과 비교한다. 마지막으로 4장에서는 논문을

결론 짓고 향후 연구 방향에 대해 논의한다.

II. 회전형 동물검출 신경망

2.1 신경망 구조

제안하는 동물검출 신경망은 학습부와 추론부로 구성된다. 먼저 동물검출 신경망의 학습부는 그림 1과 같이 동물검출모델과 부호화기, 그리고 학습기로 구성된다. 제안기술은 보편적 객체검출 방법과 달리 동물검출에 특화된 어노테이션 형식을 지닌다. 어노테이션 형식의 세부속성은 표 1과 같다.

회전형 검출상자는 그림 2에 나타난 것과 같이, 중심점을 기준으로 동물의 몸통과 일치한 방향으로 회전한 사각형을 이용하여 동물의 몸통 영역을 정의한다. 즉 동물의 회전형 검출상자는 중심점, (x_c, y_c) , 회전각, θ , 영역의 폭과 높이 (w, h) 5개의 파라미터로 표현한다.

동물검출모델은 영상으로부터 특징을 추출하는 ‘백본’, 유의한 중간정보를 추출하는 ‘넥’, 그리고 정의된 동물검출 정보를 추출하는 ‘헤드’의 세 단계로 나뉜다. 이를 수식으로 표현하면 (1)과 같다.

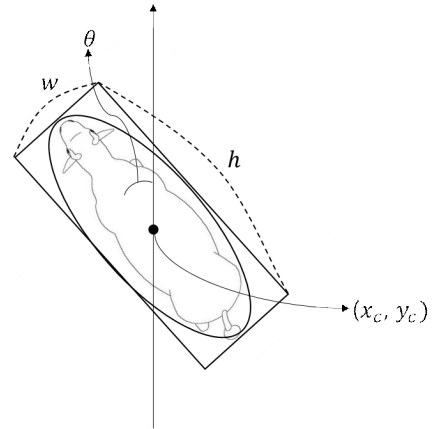


그림 2. 동물에 표시된 회전형 경계상자
Fig. 2. An illustrate of the rotated bounding box overlay to the animal

$$M(I) = H(N(B(I))) \quad (1)$$

여기서 $M(I)$ 는 RGB 이미지, I 를 입력으로 하는 동물 검출신경망 모델이며, H, N, B 는 각각 모델의 헤드, 넥, 백본을 의미한다.

그림 3은 백본, 넥, 헤드를 포함하는 동물검출모델의 전체 흐름도를 나타낸다. 이미지 모델로도 불리는 백본은 심층신경망 기반의 영상분석 및 처리 방법에서 공통으로 사용되는 구성 요소이다. 백본은 주로 2D 컨볼루션 층의 누적 형태를 띠며, 그 효율을 개선하기 위해 다양한 신경망 구조를 지니도록 개선됐다. 다양한 구조의 백본들은 공통적으로 영상을 입력받아 중간정보를 추출하는 역할을 한다. 본 논문의 동물검출모델은 ResNet^[11], DarkNet^[12], EfficientNet^[13] 등 현재까지 제시된 다양한 백본 구조를 채택할 수 있으며, 이후 개발될 새로운 구조의 백본 역시 활용할 수 있다.

다음으로 ‘넥’은 보편적 객체검출기를 구성하는 하위 신경망으로써 백본의 층별 중간정보를 취합하여 해석하는 역할을 한다. 각 층별로 해석하는 영상의 해상도가 다르므로 원/근, 동물의 체형별로 다양한 크기의 동물을 효과적으로 검출하기 위한 층별 중간정보를 헤드에 제시한다. 앞에 쓰인 백본의 형태에 따라 구체적인 구성을 달리할 수 있다. 본 논문의 동물검출모델에서는 Feature Pyramid Networks (FPN)^[14], bi-FPN^[15] 등 현재까지 제시된 다양한 넥 구조를 채택할 수 있으며, 이후 개발될 새로운 구조의 넥 역시 활용할 수 있다. 넥을 구성하는 구체적인 신경망의 층수와 층별 하이퍼파라미터는 백본의 형태에 따라 달라진다.

마지막으로 ‘헤드’는 넥에서 얻은 중간정보를 입력

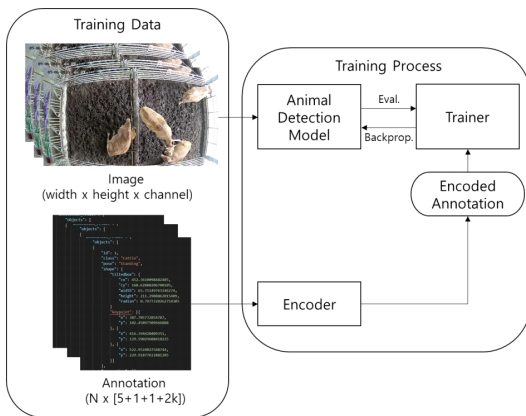


그림 1. 제안하는 회전형 동물검출 신경망의 학습절차
Fig. 1. Training process of the proposed rotated animal detection neural network

표 1. 제안하는 회전형 동물검출 신경망의 어노테이션 형식
Table 1. Annotation format of the proposed rotated animal detection neural network

Property	Format	Range
rotated bounding box	$R = \{x_c, y_c, w, h, \theta\}$	$x_c, y_c, w, h =$ [image size] $\theta = [-\pi/2 \sim \pi/2]$
class	c	Class Index, Larger than 0

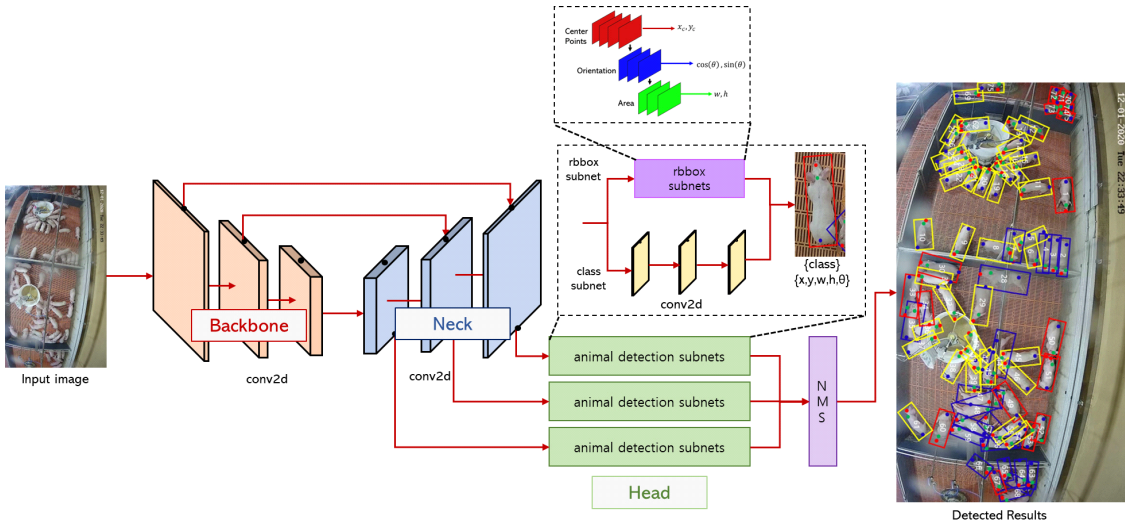


그림 3. 제안하는 동물 검출 신경망 모델의 흐름도
 Fig. 3. Flowchart of the proposed animal detection neural network model

받아 동물검출정보를 출력하는 신경망이다. 네의 각 층별 중간정보를 각각 입력받아 각 층별로 인지도된 동물검출정보를 출력한다. 보편적 객체검출과 달리 제안 기술은 동물검출 정보를 얻어내는 데 적합하도록 구조적 특징을 지닌다. 먼저 동물의 개별 영역을 정의하는 회전형 경계상자를 얻어내는 신경망은 ‘계단식 다중경로 심층 컨볼루션망’ (cascaded multi-lane deep convolutional networks) 로 구성된다. 계단식 다중경로 심층 컨볼루션망은 주어진 동물 사진에 대해서 회전형 경계상자를 찾기 위한 인과 순서를 따라 구성되었다. 영상속 하나의 동물검출정보를 정의하기 위해 다음의 순서를 따른다. a) (점) 중심점 및 주요 지점을 표기한다. b) (선) 중심점과 한 개 이상 지점을 가로지르는 접선을 긋는다. c) (면) 접선이 가운데 관통하는 영역을 지정한다. 계단식 다중경로 심층 컨볼루션망은 상기 인과순서에 따라 정보가 전달되어 각 정보를 출력한다. 즉 첫 번째 경로는 중심점을, 두 번째 경로는 접선의 방향, θ , 를, 세 번째 경로는 접선과 중심점을 포함하는 영역의 폭과 높이를 출력한다. 객체 클래스는 각각 별도의 심층 컨볼루션망 경로를 통해 얻어진다.

입력 이미지에 대한 동물검출모델, $M(I)$,의 출력은 부호화된 동물검출 정보로써 (2)와 같이 표현된다.

$$E(A) = M(I) \quad (2)$$

여기서 $A = \{R, c\}$ 는 회전형 경계상자와 클래스로 구성된 출력 벡터를 의미하고, $E(A)$ 는 부호화된 출력

벡터를 의미한다.

2.2 회전상자 부호화

제안하는 동물검출신경망 학습을 위한 동물검출정보, A ,는 학습용 이미지, I ,와 짝을 이루어 준비된다. 이때 준비 과정은 사람이 이미지를 보고 적절한 어노테이션도구¹⁶⁾를 사용하여 검출정보를 직접 입력해 주거나, 기존에 개발이 완료된 제안하는 동물검출신경망을 활용하여 자동으로 입력한 후 사람이 수정/보완하는 방법을 취한다.

동물검출모델이 입력 이미지 내 고정된 영역별로 출력값을 지니도록 출력 형태를 정의한다. 즉, 동물검출정보 부호화 과정, $E(A)$,는 다음의 절차에 따라 수행된다. 먼저 각 헤드마다 a 개의 사전 지정된 관심영역, \hat{R}_a 을 지정한다. 여기서 a 는 헤드별 처리 영역 \times 크기종류 \times 각도종류 \times 상자 비율로 정의한다. 다음으로 A 에 기입된 k 번째 동물검출영역, R_k , 와 관심영역간 겹침의 정도를 계산한다.

$$o_{a,k} = IoU(\hat{R}_a, R_k) \quad (3)$$

여기서 $IoU(x,y)$ 는 두 회전상자간 겹침정도를 계산한다¹⁷⁾. 다음으로 (4)를 통해 각 관심영역별로 가장 겹침정도가 높은 동물영역, $R_{k'}$ 을 선택한다.

$$k' = \operatorname{argmax}_k(o_{a,k}) \quad (4)$$

마지막으로 관심영역과 이에 해당하는 동물영역간 부호화를 수행한다.

$$E(A) = \nabla_a [e_R(R_k, \hat{R}_a), c_a] \quad (5)$$

여기서 $e_R(R, \hat{R})$ 은 다음의 처리 과정을 지닌다.

$$[\bar{x}_c, \bar{y}_c, \bar{w}, \bar{h}, \sin(\theta), \cos(\theta)] = e_R(R, \hat{R}) \quad (6)$$

이때 R 과 \hat{R} 은 각각

$$R = [x_c, y_c, w, h, \theta], \quad \hat{R} = [\hat{x}_c, \hat{y}_c, \hat{w}, \hat{h}, \hat{\theta}] \quad (7)$$

이고, $\bar{x}_c, \bar{y}_c, \bar{w}, \bar{h}$ 는 각각 다음과 같다.

$$\bar{x}_c = (x_c - \hat{x}_c) / \hat{w} \quad (8)$$

$$\bar{y}_c = (y_c - \hat{y}_c) / \hat{h} \quad (9)$$

$$\bar{w} = \log(w / \hat{w}) \quad (10)$$

$$\bar{h} = \log(h / \hat{h}) \quad (11)$$

2.3 역전파 학습

역전파 학습기는 동물검출모델의 학습을 위해서 부호화된동물검출정보, $E(A)$ 와 이의 추정치 간의 손실값을 계산하고 이러한 손실값이 작아지도록 신경망 파라미터를 업데이트하는 과정을 반복한다. 회전경계상자의 손실값을 계산할 때 L1 또는 L2 손실 함수를 사용하고^[8], c 의 손실값은 Binary Cross Entropy Loss^[18], 또는 Focal Loss^[19] 등의 판별 손실 함수를 사용한다. 학습은 손실 총합이 목표 수치 이하로 떨어질 때까지 반복한다.

2.4 신경망 추론

제안하는 동물검출신경망의 추론부는 그림 4와 같이 학습된 동물검출모델과 동물검출정보 복호화기로 구성된다. 추론을 위한 입력 이미지, \tilde{I} , 에서 다음 수식에 따라 부호화된 동물검출정보를 얻을 수 있다.

$$E(\tilde{A}) = M(\tilde{I}) \quad (12)$$

$E(\tilde{A})$ 로부터 최종 동물검출 출력값, \tilde{A} 를 얻기 위해서 동물검출정보 복호화 과정을 거친다. $E(\tilde{A})$ 는 \tilde{I} 에 대한 부호화된 동물검출영역, \tilde{R}_a , 및 클래스, \tilde{c}_a 로 표현

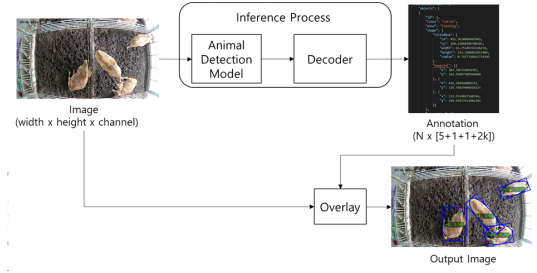


그림 4. 제안하는 회전형 동물검출 신경망의 추론 절차
Fig. 4. Inference process of the proposed rotated animal detection neural

된다.

$$E(\tilde{A}) = \nabla_a [\tilde{R}_a, \tilde{c}_a] \quad (13)$$

이때 \tilde{c}_a 가 문턱값보다 높은 후보 값들만 남기는데 수식으로 표현하면 다음과 같다.

$$a' = a \text{ if } \tilde{c}_a > thr \quad (14)$$

여기서 thr 은 검출문턱값이다. 다음으로 $d_R(\tilde{R}_a, \hat{R}_a)$ 을 수행하여 복호화된 동물검출영역, \bar{R} ,을 얻는다. 이때 $\bar{R} = d_R(\tilde{R}_a, \hat{R}_a)$ 의 처리과정은 다음과 같다.

$$\begin{aligned} \tilde{R} &= [\tilde{x}_c, \tilde{y}_c, \tilde{w}, \tilde{h}, \sin(\tilde{\theta}), \cos(\tilde{\theta})], \\ \bar{R} &= [\bar{x}_c, \bar{y}_c, \bar{w}, \bar{h}, \theta] \end{aligned} \quad (15)$$

이고 $\bar{x}_c, \bar{y}_c, \bar{w}, \bar{h}, \sin(\tilde{\theta}), \cos(\tilde{\theta})$ 는 각각

$$\bar{x}_c = \tilde{x}_c \hat{w} + \hat{x}_c \quad (16)$$

$$\bar{y}_c = \tilde{y}_c \hat{h} + \hat{y}_c \quad (17)$$

$$\bar{w} = \exp(\tilde{w}) \hat{w} \quad (18)$$

$$\bar{h} = \exp(\tilde{h}) \hat{h} \quad (19)$$

$$\bar{\theta} = \text{atan}(\sin(\tilde{\theta}) / \cos(\tilde{\theta})) \quad (20)$$

와 같이 얻어진다.

마지막으로 복호화 된 \bar{R} 는 하나의 동물에 대하여 다수의 중첩된 경계상자로 표현할 수 있다. 이를 하나의 결과로 표현하기 위해 non-max suppression^[20] 알고리

즘을 통해 중첩된 경계상자를 없앤다.

III. 실험

본 절에서는 제안된 회전형 동물 검출 신경망의 성능을 평가하기 위한 실험환경 설정 및 평가 결과에 대해 다룬다.

3.1 데이터 준비

회전형 동물 검출 신경망의 성능을 평가하기 위해 7개의 실제 농장과 1개의 가상 농장에서 26,208장의 돼지 이미지를 수집하였다. 실제 농장의 이미지는 디지털 카메라를 사용하여 1920x1080의 full-HD 해상도로 취득하였다. 데이터 수집에 협조한 농장들은 국내의 자돈생산농장 2곳, 육성비육 농장 5곳이며 각 농장들의 사육 품종은 3원교잡종 (Yorkshire, Landrace, Duroc, YLD)이었다. 이미지 데이터 수집을 위해 각 농장마다 1개 이상의 IP카메라를 돈사 내에 설치하였다. 카메라 설치 지점은 각각 약 20일령부터 180일령까지 다양한 성장단계의 돼지가 살고 있는 군사 공간을 선택하여 데

이터셋에 다양한 크기의 돼지가 포함되도록 하였다. 영상 수집 기간은 일주일이며, 각 카메라 별로 녹화한 영상들 중에서 중복성이 낮고 돼지의 등장 빈도가 높은 이미지들을 추출하였다. 실제 이미지 데이터의 어노테이션은 숙련된 사람 작업자 3인이 웹기반 사용자 인터페이스를 통해 이미지를 눈으로 보며 마우스를 이용하여 회전 경계상자와 클래스를 수동으로 기록하는 방식으로 진행하였다.

가상 농장의 경우 Unity Engine²¹¹을 이용하여 실제 농장을 모사한 3D 가상공간의 농장 내에 약 6kg~120kg에 해당하는 성장단계의 3D 가상 돼지 객체를 배치하였다. 이때 돼지의 품종별 외형과 색깔이 다양한 점을 가상데이터에 반영하기 위해 요크셔, 버크셔, 듀록, 흑돼지 총 4종의 3D 가상 돼지 객체 모델을 만들었다. 이렇게 구성된 가상 농장 공간에서 실제 농장에 설치된 카메라와 유사한 화각이 표현되는 지점을 선정 후, 스냅샷을 촬영하는 방식으로 가상 이미지를 추출하였다. 가상 이미지 데이터의 어노테이션은 3D 객체별 좌표 정보를 활용하여 회전 경계상자와 클래스를 자동으로 기록하였다.

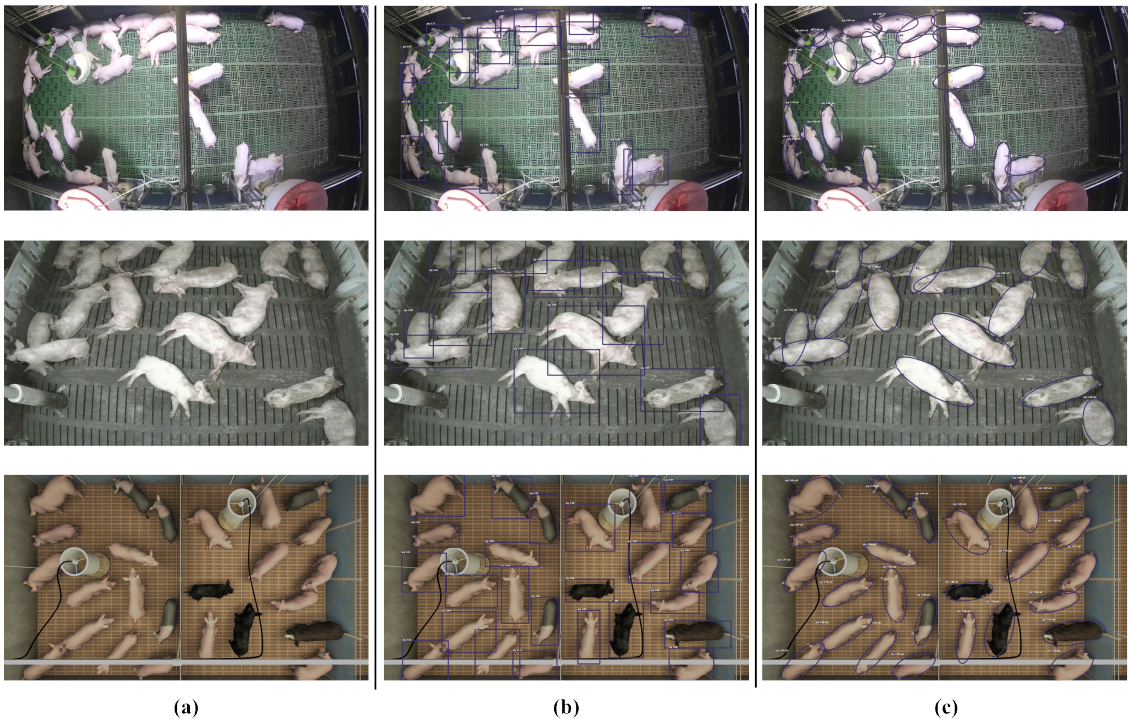


그림 5. 양돈 영상에서의 양돈 검출 결과. (a) 입력 이미지, (b) 기존의 직사각형 경계상자 검출결과, (c) 제안된 회전형 경계상자 검출결과

Fig. 5. Pig detection results for pig farming images. (a) Input image, (b) Detection results by the conventional rectangular bounding box detection, (c) Detection results by the proposed rotated bounding box detection

표 2. 동물검출 방법들의 검출 성능 지표 측정 결과
Table 2. Results on object detection performance metrics of the animal detection methods

Method	FP	FN	TP	TN	Precision	Recall	Accuracy	F1-score	mAP50	mAP50:95
YOLOv7-tiny with rectangular bounding box	0.1368	0.0158	0.8632	0.9842	0.8632	0.9820	0.9237	0.9188	0.8743	0.4018
Proposed method with rotated bounding box (Relative changes)	0.0557 (-59.28%)	0.0066 (-58.23%)	0.9443 (9.40%)	0.9934 (0.93%)	0.9443 (9.40%)	0.9931 (1.12%)	0.9689 (4.89%)	0.9681 (5.36%)	0.9534 (9.05%)	0.6304 (56.89%)

전체 이미지를 학습 및 평가하기 위해 훈련 세트와 평가 세트도 나누었다. 훈련 세트는 24,875개의 이미지로, 평가 세트는 1,333개의 이미지로 구성되도록 나누었다.

3.2 동물검출 신경망 준비

제안하는 회전형 동물검출 신경망의 성능을 비교하기 위해 사각형 경계상자를 출력하는 기존의 객체검출 신경망과 제안 방법을 준비 하였다. 구체적으로 비교를 위해 최근에 공개되어 처리 속도와 검출 성능이 우수한 기존의 객체 검출 신경망, YOLOv7^[22], 을 사용하였다. 본 실험에서는 YOLOv7-tiny 설정을 활용하였으며 하이퍼파라미터와 데이터증강설정 등은 모두 공개된 설정을 따랐다^[23]. 기존방법과 제안방법 간의 공정한 성능 비교를 위해 제안방법의 백본 및 neck의 구조와 하이퍼파라미터는 모두 YOLOv7-tiny의 것을 동일하게 사용하였다. 마지막으로 제안방법과 달리 기존방법은 회전경계상자를 처리할 수 없으므로 준비된 데이터셋 내 회전경계상자의 외접상자를 계산하고 이 외접상자를 기존방법의 학습 및 평가에 활용하였다.

3.3 성능 측정

제안한 회전형 동물검출 신경망의 성능을 평가하기 위해 1,333장의 이미지로 구성된 평가용 데이터셋을 활용하였다. 우선 제안 방법의 검출 성능을 분석하기 위해 객체별 분류 성능 평가 지표^[24]를 측정하였다. 이를 위해 기존방법과 제안방법 각각에 대한 모든 평가 이미지 내 포함된 돼지의 개체별 검출 결과를 집계하였다. 이때 모든 방법의 검출 민감도 파라미터는 0.9로, IoU (IoU)는 0.5로 설정하였다. 검출 결과 집계치로부터 true positive (TP), false positive (FP), true negative (TN), false negative (FN) 각각의 비율을 측정하였다. 이 값들로부터 정확도 (precision), 검출률 (recall), 정확성 (accuracy), 그리고 F1-score를 계산하여 검출 성능을 전반적으로 비교하였다. 추가로 객체 검출 성능을 비교하는데 보편적으로 활용되는 평가지표인 mean average

precision (mAP)^[17] 역시 비교 측정하였다. 기존의 YOLOv7-tiny와 제안한 회전형 동물검출 신경망의 평가지표들은 표 2와 같다.

표 2에 나타난 것과 같이 제안된 방법은 기존 방법대비 높은 TP와 낮은 FP를 보이며 정확도, 검출률, 정확성, 그리고 F1-score에서 각각 9.40%, 1.12%, 4.89%, 그리고 5.36% 개선된 결과를 보였다. 또한 mAP 분석 측면에서 제안된 방법은 0.9534의 mAP50과 0.6304의 mAP50:95를 얻으며 기존 방법보다 각각 9.04%, 56.89% 높은 성능을 보여주었다. 이러한 결과는 제안한 방법이 가축 동물을 검출하는 데 기존의 직사각형 경계 상자 방식보다 검출성능이 우수함을 보여준다. 그림 5는 평가용 데이터셋의 일부인 이미지로부터 기존방법과 제안방법으로 처리한 결과를 보여준다. 그림에서 확인할 수 있듯이, 기존의 직사각형 경계상자 방식은 가축동물이 밀집한 상황에서 각 객체를 표현하는 직사각형 경계상자간 겹침이 높아 부정확한 결과를 보인다. 반면, 제안된 방법은 회전경계상자로 동물을 검출하여 여러 마리가 밀집한 상황에서도 각 가축동물의 영역을 정확히 검출한 것을 볼 수 있다.

다음으로 제안하는 방법과 기존 방법 간의 연산 시간도 비교하였다. 연산 시간의 비교를 위해 임베디드용 엡지컴퓨팅 보드로 널리 쓰이는 NVIDIA사의 Jetson Nano 2GB Developer Kit^[25]을 활용하였다. 제안방법과 기존방법을 모두 준비된 엡지컴퓨팅 보드에 구현하였으며, 이때 두 방법 모두 최적의 성능을 발휘할 수 있도록 TensorRT^[26]를 이용하여 floating point 16 양자화 및 최적화를 진행하였다. 표 3은 기존방법과 제안방법의 GFLOPS (Giga Floating point Operations Per Second)와 FPS (Frames Per Second)를 보여준다.

표 3과 같이 제안방법은 기존의 YOLOv7-tiny 대비 0.9 증가한 GFLOPS를 보이며, 동일한 HW 환경에서 거의 유사한 수준의 FPS를 보여주었다. 이는 제안하는 회전형 동물 검출 신경망이 요구하는 HW 자원이 고도로 최적화 된 YOLOv7-tiny에 비교하여 큰 차이가 없이 효율적임을 의미한다.

표 3. 동물검출 방법들의 연산량 및 처리속도 비교
Table 3. Comparison of calculation amount and processing speed of animal detection methods

Method	GFLOPS	FPS
YOLOv7-tiny with rectangular bounding box	12.1	29.89
Proposed method with rotated bounding box (Relative changes)	13.0 (7.44%)	29.87 (-0.07%)

IV. 결 론

본 논문에서는 축산동물 관리를 위해 기존의 직사각형 출력 방식보다 정확하게 각 동물의 폭, 길이 및 회전 방향을 감지하는 회전형 동물 검출 신경망을 제안하였다. 제안된 방법을 평가하기 위해, 7개 농장과 1개 가상 농장에서 26,208개의 돼지 사육장 이미지 데이터셋을 수집하고, 학습에서 배제된 1,333개의 평가용 이미지를 사용하여 실험을 진행하였다. 실험 결과, 제안된 방법은 기존방법보다 9.04% 높은 0.9534의 mAP50 값을 달성하였다. 또한 제안된 방법은 13.0 GFLOPS의 계산량과 임베디드 장치에서 29.8FPS의 속도를 보였다. 이러한 실험 결과에 따라 제안된 방법이 개체 간 밀집이 일반적인 인터넷 연결이 제한적인 실제 축사 환경에서 다수의 동물을 실시간으로 검출하는데 적합한 기술임을 입증하였다. 따라서 제안된 방법은 향후 한우, 양돈, 양계 등 주요 축산동물의 영상 기반 비접촉 생체정보 측정에 활용될 것으로 기대한다. 축종별 비접촉 생체정보측정을 위해서는 제안된 방법과 개체 추적, 키포인트 검출, 행동분석 등의 연관 기술과의 융합에 관한 후속 연구가 요구된다.

References

[1] P. J. Gerber, H. Steinfeld, B. Henderson, A. Mottet, C. Opio, J. Dijkman, and G. Tempio, *Tackling climate change through livestock: A global assessment of emissions and mitigation opportunities*, Food and Agriculture Organization of the United Nations (FAO), 2013.

[2] C. J. Rutten, A. G. J. Velthuis, W. Steeneveld, and H. Hogeveen, "Sensors to support health management on dairy farms," *J. Dairy Sci.*, vol. 96, no. 4, pp. 1928-1952, 2013.

(<https://doi.org/10.3168/jds.2012-6107>)

[3] K. Schirmann, N. Chapinal, D. M. Weary, W. Heuwieser, and M. A. G. von Keyserlingk, "Rumination and its relationship to feeding and lying behavior in Holstein dairy cows," *J. Dairy Sci.*, vol. 95, no. 6, pp. 3212-3217, 2012.
(<https://doi.org/10.3168/jds.2011-4741>)

[4] S. Wolfert, L. Ge, C. Verdouw, and M. J. Bogaardt, "Big data in smart farming-a review," *Agric. Syst.*, vol. 153, pp. 69-80, 2017.
(<https://doi.org/10.1016/j.agry.2017.01.023>)

[5] A. Nasirahmadi, O. Hensel, S. A. Edwards, and B. Sturm, "Implementation of machine vision for detecting behaviour of cattle and pigs," *Livestock Sci.*, vol. 202, pp. 25-38, 2017.
(<https://doi.org/10.1016/j.livsci.2017.05.014>)

[6] Q. Guo, Y. Sun, C. Orsini, J. E. Bolhuis, J. de Vlieg, P. Bijma, and P. H. de With, "Enhanced camera-based individual pig detection and tracking for smart pig farms," *Comput. and Electr. in Agric.*, vol. 211, 108009, 2023.
(<https://doi.org/10.1016/j.compag.2023.108009>)

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. CVPR*, pp. 779-788, 2016.
(<https://doi.org/10.1109/CVPR.2016.91>)

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Advances in NIPS*, pp. 91-99, 2015.
(<https://doi.ieeecomputersociety.org/10.1109/TPAMI.2016.2577031>)

[9] Y. Wei, H. Zhang, C. Gong, D. Wang, M. Ye, and Y. Jia, "Study of pose estimation based on spatio-temporal characteristics of cow skeleton," *Agric.* vol. 13, no. 8, 1535, 2023.
(<https://doi.org/10.3390/agriculture13081535>)

[10] T. Nath, A. Mathis, A. C. Chen, A. Patel, M. Bethge, and M. W. Mathis, "Using DeepLabCut for 3D markerless pose

- estimation across species and behaviors,” *Nature protocols*, vol. 14, no. 7, pp. 2152-2176, 2019.
(<https://doi.org/10.1038/s41596-019-0176-0>)
- [11] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. CVPR*, pp. 770-778, 2016.
(<https://doi.org/10.1109/CVPR.2016.90>)
- [12] J. Redmon, “*Darknet: Open source neural networks in C*,” 2016.
(<http://pjreddie.com>)
- [13] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in *Proc. 36th ICML*, pp. 6105-6114, 2019.
(<https://doi.org/10.48550/arXiv.1905.11946>)
- [14] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature pyramid networks for object detection,” in *Proc. IEEE Conf. CVPR*, pp. 2117-2125, 2017.
(<https://doi.org/10.1109/CVPR.2017.106>)
- [15] M. Tan, R. Pang, and Q. V. Le, “Efficientdet: Scalable and efficient object detection,” in *Proc. IEEE/CVF Conf. CVPR*, pp. 10781-10790, 2020.
(<https://doi.ieeecomputersociety.org/10.1109/CVPR42600.2020.01079>)
- [16] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, “LabelMe: A database and web-based tool for image annotation,” *Int. J. Comput. Vision*, vol. 77, no. 1-3, pp. 157-173, 2008.
(<https://doi.org/10.1007/s11263-007-0090-8>)
- [17] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, “The PASCAL visual object classes (VOC) challenge,” *Int. J. Comput. Vision*, vol. 88, no. 2, pp. 303-338, 2010.
(<https://doi.org/10.1007/s11263-009-0275-4>)
- [18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
(<https://dl.acm.org/doi/book/10.5555/3086952>)
- [19] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, “Focal loss for dense object detection,” in *Proc. IEEE ICCV*, pp. 2980-2988, 2017.
(<https://doi.org/10.1109/ICCV.2017.324>)
- [20] A. Neubeck and L. Van Gool, “Efficient non-maximum suppression,” in *Proc. 18th ICPR'06*, vol. 3, pp. 850-855, 2006.
(<https://doi.org/10.1109/ICPR.2006.479>)
- [21] Unity Technologies, “*Unity*,” 2021.
(<https://unity.com/>)
- [22] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” *arXiv preprint arXiv:2207.02696*, 2022.
(<https://arxiv.org/abs/2207.02696>)
- [23] “*YOLOv7: State of the Art Object Detection Model*,” 2021.
(<https://github.com/WongKinYiu/yolov7>.)
- [24] M. Hossin and M. N. Sulaiman, “A review on evaluation metrics for data classification evaluations,” *Int. J. Data Mining & Knowledge Manag. Process*, vol. 5, no. 2, pp. 1-11, 2015.
(<https://doi.org/10.5121/ijdkp.2015.5201>)
- [25] NVIDIA Corporation, “*Jetson Nano 2GB Developer Kit*,” 2021.
(<https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-nano-2gb>)
- [26] NVIDIA Corporation, “*NVIDIA TensorRT*,” 2021.
(<https://developer.nvidia.com/tensorrt>)

전 광 명 (Kwang Myung Jeon)



2010년 2월 : 세종대학교 정보통신공학과 학사
2012년 2월 : 광주과학기술원 정보통신공학부 석사
2019년 8월 : 광주과학기술원 전자전자컴퓨터공학부 박사
2019년 8월~현재 : 인트플로우

주식회사 대표이사
<관심분야> 딥러닝, 영상처리, 음성처리, 경량화, 스마트팜

[ORCID: 0000-0003-2810-3063]

권 솔 빈 (SolBeen Kwon)



2019년 2월 : 조선대학교 컴퓨터통계학과 학사
2020년 9월~현재 : 인트플로우 주식회사 선임연구원
<관심분야> 딥러닝, 영상처리, 경량화, 엣지컴퓨팅
[ORCID:0009-0001-1459-9523]

안 형 준 (HyungJun Ahn)



2020년 2월 : 목포대학교 컴퓨터공학과 학사
2020년 1월~현재 : 인트플로우 주식회사 선임연구원
<관심분야> 딥러닝, 영상처리, 음성처리

[ORCID:0009-0001-5338-1017]

정 진 우 (Jinwoo Jung)



2011년 2월 : Pakuranga College 졸업
2012년 12월 : Unitec Institute of Technology (Management)
2021년 3월~2021년 11월 : 주식회사 꼬리 연구원
2022년 6월~현재 : 인트플로우 주식회사 연구원

<관심분야> 딥러닝, 기계학습, 데이터 전처리
[ORCID:0009-0003-6885-9874]

주 소 현 (Soheun Ju)



2019년 8월 : 전남대학교 통계학과 학사
2020년 6월~현재 : 인트플로우 주식회사 선임연구원
<관심분야> 딥러닝, 영상처리, 통계분석

[ORCID:0009-0008-1233-2084]

이 용 기 (Younggie Lee)



1995년 2월 : 전남대학교 농과대학 축산학과 학사
2011년 4월~현재 : 성일축산영농조합법인 대표이사
2018년 1월~현재 : 한국농수산대학교 현장교수
2018년 3월~현재 : 전남대학교

동물자원학부 석박사 통합과정 재학중
<관심분야> 양돈, 축산경영, 융합기술, 스마트팜
[ORCID:0009-0009-1522-6014]

류 인 철 (Inchul Ryu)



2017년 2월 : 전북대학교 기계시스템공학부 학사
2020년 7월~현재 : 인트플로우 주식회사 선임연구원
<관심분야> 딥러닝, 영상처리, 임베디드 시스템

[ORCID:0009-0002-8999-0073]