



실시간 말벌 탐지를 위한 향상된 타일링 기법

권영재, 정윤주¹, 이철희*

안동대학교 컴퓨터공학과, ¹안동대학교 SW융합교육원

Improved Image Tiling Method for Real-time Vespa Detection

Yeong Jae Kwon, Yunju Jeong¹ and Cheol Hee Lee*

Department of Computer Engineering, Andong National University
¹Institute of Software Convergence Education, Andong National University

Abstract

In order to effectively control wasps, which are the main cause of damage in beekeeping, a monitoring system that can check the appearance of wasps in real time is necessary. Convolutional Neural Networks (CNNs), currently widely used in the field of object recognition, cannot successfully detect and classify small objects such as wasps. When a wasp is photographed at a distance of 40 cm with a 4K-camera, the size of the wasp is only 2~3% of the total image. Therefore, in order to increase the recognition accuracy of small objects, we use a tiling method to detect wasps by dividing the image into 3×3. In the traditional tiling methods do not recognize objects located on the divided boundary well. In order to improve the object recognition performance of the tiling method, we propose an improved tiling method that additionally detect a partial area centered on the boundary line in the original image when there are some objects on the tile boundary line. We evaluated the performance of the proposed tiling method using the YOLOX model after producing 7,505 4K resolution images in which 5 species of wasps and 1 type of bee were randomly placed on a 3×3 tile boundary. In experiments on image data created for wasp object recognition training, the PASCAL VOC mAP of traditional tiling methods was 3.28%. However, the mAP of the proposed tiling method was 14.14%, which showed better performance of the proposed tiling technique in terms of accuracy.

Keywords

Object detection, CNN, Vespa monitoring system, YOLOX, Tiling

서 론

말벌 피해는 양봉가의 봉군 유지의 가장 큰 어려움이다 (Jung, 2012a; Kim *et al.*, 2017). 말벌은 8월 상순부터 11월 초순 사이에 피해를 주며 벌통 전체가 폐사할 수 있다고 보고된 바 있다 (Jung *et al.*, 2007, 2012b). 등검은말벌 (*V. velutina nigrithorax*)의 출현은 토착 말벌류 개체군의 구성비 변화를 일으킨다고 보고되고 있다 (Jung, 2012). 이러한 피해를 막기 위한 연구로 말벌의 음향정보를 해석하여 조기경보 시스템을 구현하거나 (Kim *et al.*, 2017), 심층 합

성곱 신경망 (DCNN, Deep Convolutional Neural Network)을 사용하여 입력 영상 내에서 말벌을 검출하는 방법이 보고되고 있다 (Lee *et al.*, 2019; Jeong *et al.*, 2020). DCNN을 사용한 객체 탐지 분야에서 목표 객체의 크기가 작은 경우 그렇지 않은 객체에 비해 2~3배의 정확도 감소가 발생하는 것으로 알려져 있다 (Kisantal *et al.*, 2019). 등검은 말벌 일벌의 경우 길이가 22~25 mm로 확인되었다 (Jung *et al.*, 2009). 그러므로 카메라에 촬영되는 영상에서 말벌의 크기는 신경망에서 요구되는 최소한의 검출 크기보다 작아, 원활한 검출을 위해 추가적인 인식 방법이 요구

된다. 작은 객체 검출 개선을 위해(Lin *et al.*, 2017)은 서로 다른 스케일의 피쳐맵(Feature map)을 생성 후 결합하는 시도가 있었다. 또한 적대적 생성 신경망(GAN)으로 입력 영상의 선명도를 올려 검출을 개선하는 방법도 제시되었다(Bai *et al.*, 2018). 그러나 실시간 처리를 위해서는 YOLO와 같이 객체의 위치와 분류가 동시에 이루어지는 일괄 검출(one stage detection)기반의 딥러닝 기술이 필요하다. Stephane Charette은 객체의 최소 입력 크기를 보장하기 위해 입력 영상을 여러 구역으로 분할 후 YOLO기반 분류기를 이용하여 대상 객체를 검출한 뒤 잘린 바운딩 박스를 보정, 구역을 결합해 하나로 만드는 타일링 기술을 제안하였다. 하지만 해당 타일링 방법에서 잘린 바운딩 박스를 올바르게 처리하기 위해서는 잘린 바운딩 박스가 동일한 객체로 분류되어야 하며 두 개가 요구되는 한계점이 있다. 따라서 본 연구에서는 위의 한계점을 개선하기 위하여 두 개의 바운딩 박스가 다르거나, 하나만 검출되는 경우 추가 검출과정을 도입하여 기존 바운딩 박스를 이어 붙이는 방식이 아닌, 새로운 바운딩 박스를 만들어 대체하는 알고리즘을 제안하였다. 제안한 방법의 효과를 검증하기 위해 경계 영역에 객체가 집중적으로 분포하는 전용 테스트 데이터를 만들어 타일링시 발생하는 상기의 문제를 직접적으로 비교하였으며 PASCAL VOC의 mAP(mean average precision)를 기준으로 기존방법과 비교하였다. 결과적으로 AP, mAP면에서 기존방법보다 우수함을 확인하고자 한다. 또한 본 연구는 대상 객체가 매우 작은 말벌 모니터링 분야에서 필수적인 전처리 방법으로 개선된 타일링 기법을 제시한다.

재료 및 방법

1. 타일링 알고리즘

심층 합성곱 신경망을 사용한 영상 인식, 분류에서 입력 영상의 크기는 제한된다. 예를 들어 YOLOv3(Redmon and Farhadi, 2018)는 가로 및 세로 각각 416화소의 해상도를 요구하고 YOLOv4(Bochkovskiy *et al.*, 2020)는 가로 및 세로 각각 512화소의 크기를 기본값으로 요구한다. 즉, 해상도가 커질수록 정보량은 증가하지만, 연산량이 기하급수적으로 증가하여 실시간 처리가 불가능하다. 따라서 제한된 크기의 영상에서 인식 성능을 향상하기 위해서는 영상에서 대상 객체가 차지하는 면적이 커야 한다. 그

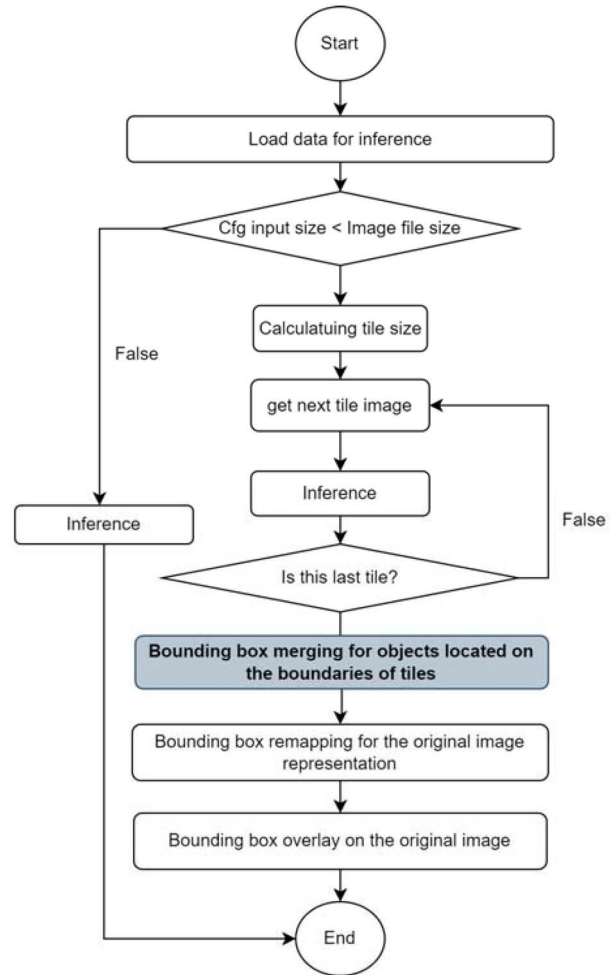


Fig. 1. Flow chart for Tiling.

러나 말벌과 같이 작은 객체의 경우 입력 영상에서 차지하는 비율이 상대적으로 작아 가장 우수한 성능의 심층 합성곱 신경망을 적용하더라도 객체 인식에 어려움이 있다. 따라서 한 장의 영상을 분할하여 여러 장의 타일 형태로 입력하여 인식 성능을 향상하는 전처리 방법이 요구된다. Fig. 1은 타일링의 전체적인 기능을 순서도로 표현한 것이다. 먼저 영상의 크기를 고려하여 생성할 타일의 수를 결정하고 각 타일 영상에 대하여 추론하며, 인접한 타일 영상에서 중복되는 객체의 경우 바운딩 박스에 대한 재처리를 통해 경계 영역에 중복되는 객체에 대한 새로운 바운딩 박스 처리가 필수적이다.

Fig. 2는 타일링에 대한 세부적인 설명을 위해 추가한 그림이다. Fig. 2(a)를 입력 영상으로 하고 타일 개수를 9개(3×3)로 설정했을 때, 검출을 수행하게 될 영역은 Fig. 2(b)와 같다. 숫자의 순서대로 마지막 9번 타일까지 9번

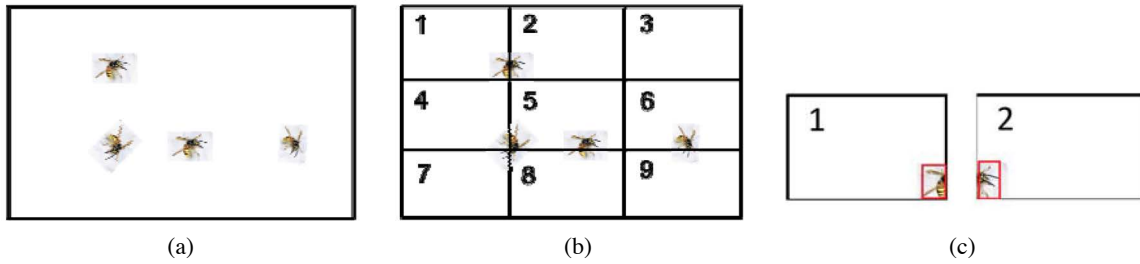


Fig. 2. Visualization of tiling process. (a) Input image, (b) 9 tile images (3×3 case), (c) Example of object detection for tile1 and tile2.

검출을 수행하게 되는 것이다. 각 타일 영상별로 검출을 마친 후 Fig. 2(c)와 같이 경계영역에 존재하는 하나의 객체를 절반씩 두 번 검출해 두 개의 바운딩 박스가 생성된다. 이것을 올바르게 처리하기 위한 단계가 필요하며, 그것이 다음으로 이어지는 바운딩 박스 결합 과정이다.

Stephane의 타일링 알고리즘은 Fig. 3의 조건문을 따른다. 이 경우 Fig. 4의 (a)처럼 두 바운딩 박스의 식별 결과가 일치하고, 두 박스가 타일링 경계선을 사이에 두고 연속적으로 존재할 경우, 두 박스의 좌표 정보를 조합하여 박스를 이어 붙인다. 그 결과가 Fig. 4의 (b)이다.

Stephane의 처리 방법은 Fig. 5의 (b)를 개선할 수 없다. 왜냐하면 두 개의 검출 결과 중 어느 쪽이 올바른 것인지 판단할 방법이 없다. 추가로 Fig. 6의 (a)처럼 타일링의 경계에 두 객체(ggoma와 jangsu)가 인접하여 존재할 수도 있으므로 두 박스를 하나로 합치는 것만으로는 모든 문제에 대응할 수 없다. 따라서 제안하는 방식에서는 우선 Fig. 5의 (b)를 찾기 위해 클래스 일치 여부를 확인하지 않고 타일링 경계선에 놓인 두 박스의 연속성 여부만 파악한다. 그러므로 Fig. 7의 첫 번째 조건문은 Fig. 4의 조건문과 달리 두 클래스가 같은지 확인하지 않는다. 이렇게 Fig. 5의 (b) 케이스를 찾아낸다. 그 후에 두 박스의 합산된 면적만큼 새로운 검출 영역을 지정한다. 이 영역은 항상 타일링의 경계선을 포함하기 때문에 한 객체에 두 개의 박스가 생성되지 않는다. 새로운 검출로 올바르게 생성된 박스를 원래의 박스 대신 영상에 나타내면 Fig. 6의 (a)와 (b)처럼 두 가지 경우와 무관하게 정상적으로 처리할 수 있다. Stephane의 처리 방법은 Fig. 5의 (c)도 해결될 수 없는데 Fig. 4의 조건문은 두 개의 박스를 대조하면서 진행되기 때문이다. 제안하는 방식에서는 Fig. 5의 (c)를 해결하기 위해 Fig. 7의 두 번째 조건문을 추가하였다. 해당 조건문에서는 바운딩 박스 처리 과정을 겪지 않았으며, 동시에 경계선에 인접한 박스를 찾는다. 그 후 똑같이 박스의 면

적을 추론 영역으로 사용하는데 이번에는 박스가 오직 1개만 존재하므로 타일링의 경계선과 수직인 방향으로 박스의 크기를 확장하여 검출 영역을 지정한다. 마찬가지로 새로운 검출로 올바르게 생성된 박스를 본래의 박스 대신 영상에 나타내면 Fig. 5의 세 가지 모든 경우를 처리할 수 있게 된다.

2. 학습 데이터의 구성

학습에는 한국에 분포하는 말벌 종(Choi *et al.*, 2013)을 고려하여 본 연구에서는 5종의 말벌(등검은말벌: *V. velutina nigrithorax*, 장수말벌: *V. mandarinia*, 꼬마장수말벌: *V. ducalis*, 털보말벌: *V. similima*, 말벌: *V. crabro*)과 양봉꿀벌(*Apis mellifera*)에 대해 총 7,505장의 영상을 사용했다. 모든 영상은 말벌 전문가로부터 영상별로 종의 분류에 대한 사전 검증을 받았다. 학습용 데이터의 작성을 위해 Table 1의 영상에 대해 각 객체의 위치 및 종류에 대하여 바운딩 박스(bounding box)처리 및 레이블링했으며 Fig. 8는 레이블링이 적용된 학습 데이터의 예시이다.

3. 테스트 데이터 생성

타일링 방법에 따른 성능 비교를 위해서는 경계 영역에 집중적으로 배치된 말벌 영상이 필요하다. 따라서 본 실험에서는 동일한 객체 검출 환경에서 두 개의 타일링 알고리즘 성능만 따로 비교하고자 Fig. 9의 절차를 따라 전용 테스트 데이터를 생성했다. 실험에 사용된 카메라의 해상도를 고려하여 4K 해상도(3840×2160)의 검은색 배경을 선언하고 Table 1의 test 데이터를 모두 불러와 같은 벌이 연속하지 않는 조건으로 불러온 영상의 순서를 무작위로 섞는다. 이후 영상을 YOLOX(Ge *et al.*, 2021)로 검출하고 검출된 바운딩 박스를 기준으로 말벌 영상을 잘라낸다. 잘라낸 영상을 Fig. 5의 잘못된 사례가 발생하는 타

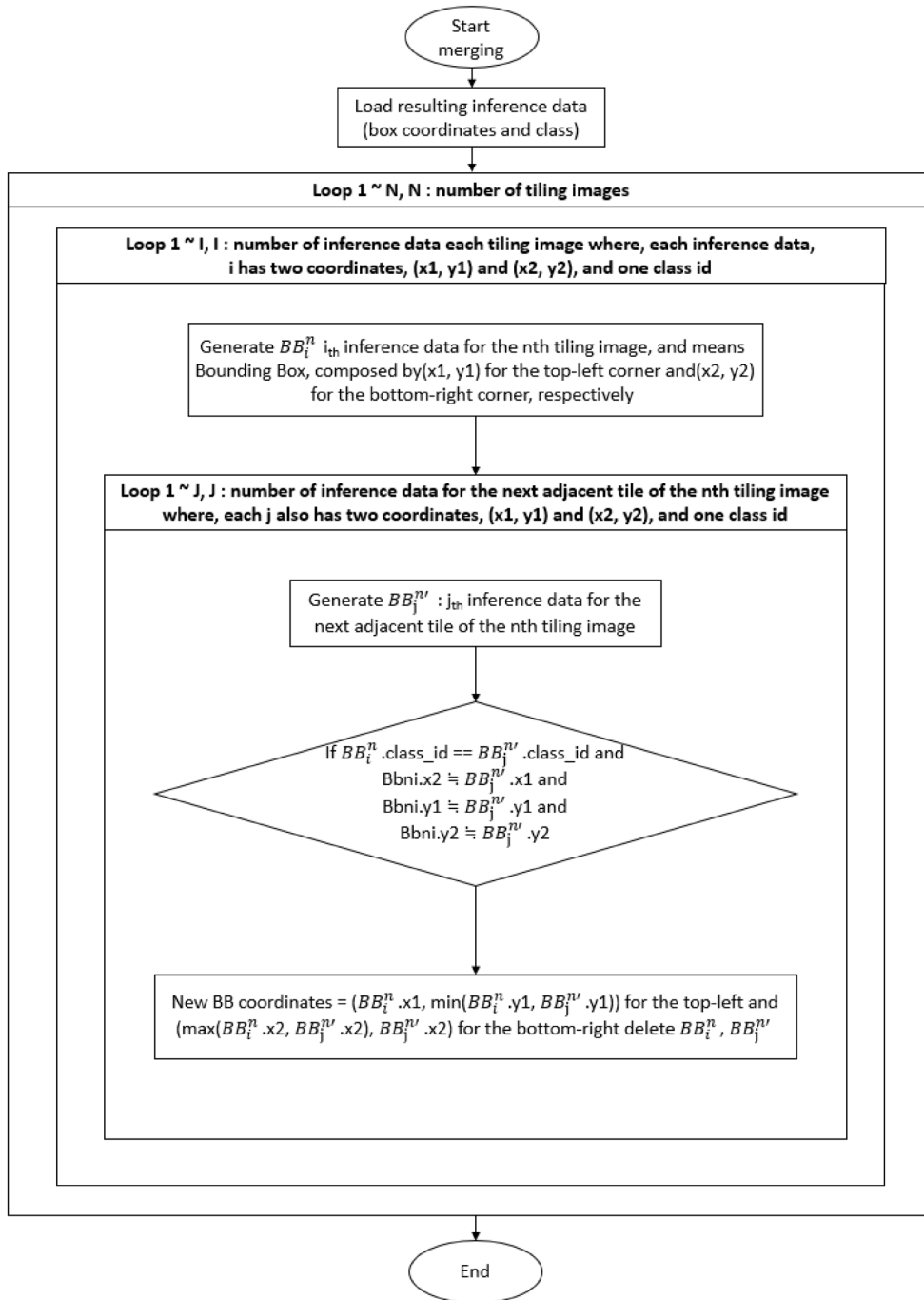


Fig. 3. Stephane's bounding box merging algorithm.

일링 경계선 주변 영역에 풀고루 배치한다. 이때 배치한 위치 정보를 txt 파일로 작성하여 영상에 대한 라벨링 데이터를 생성한다. 이 과정을 타일 영상의 경계가 가득 찰 때까지 반복한다. 테스트 영상에서 타일링 크기는 3×3으로 고정했다. 따라서 3×3의 경계 영역에 다수의 말벌 또는 꿀벌 이미지가 배치되어 있다. 이 방법으로 총 150장의

테스트 영상을 제작해 타일링에 따른 mAP를 측정하기 위한 데이터로 사용했다.

Table 2는 추론에 따른 경우의 수를 정리한 것이다. Precision은 Positive로 추론한 것 중에서 TP(정확한 추론)의 비중을 의미하고, Recall은 Ground truth가 Positive인 것 가운데 Positive로 추론한 비중을 의미한다. 그리고 객



Fig. 4. Stepane C. method. (a) Before merging, (b) After merging.

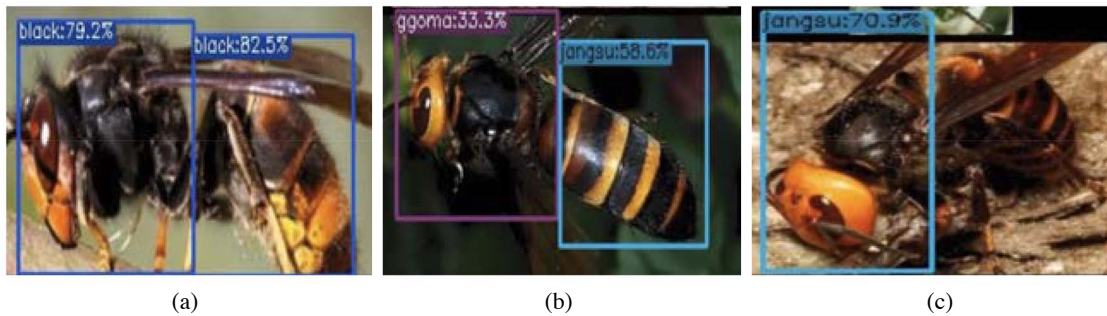


Fig. 5. Errors in detection process using tiling. (a) Two bounding boxes for an object, (b) Two bounding boxes and two classes for an object, (c) One bounding box with half size.



Fig. 6. VarIOUs detection results on the boundary area. (a) Correct detection for two objects, included in different classes, (b) Incorrect detection from one object.

체 탐지일 경우 TP를 정의할 때 추론한 객체의 위치(바운딩 박스)와 실제 객체의 위치 사이의 중첩비율을 의미하는 IOU가 포함되며 이는 Pascal VOC, COCO Dataset에 따라 다양한 기준이 존재한다. 본 논문에서는 Pascal VOC에

Table 1. Training data set.

Scientific name	Label	Train	Validation	Test
<i>Apis mellifera</i>	Apis	1115	150	150
<i>V. ducalis</i>	Ggoma	1285	150	150
<i>V. similima</i>	Simil	686	150	150
<i>V. velutina nigrithorax</i>	Balck	941	150	150
<i>V. mandarinia</i>	Jangsu	762	150	150
<i>V. crabro</i>	Crabro	916	150	150

Table 2. Type 1, Type 2 Errors in Deep Learning Object Detection.

Ground truth	Predict result	
	Positive	Negative
Positive	TP (True Positive)	FN (False Negative)
Negative	FP (False Positive)	TN (True Negative)

서 정한 IOU 기준에 따라, 각각의 클래스별 추론 결과를 이용해 x축(recall)과 y축(precision)에 대하여 AP(Average Precision)의 면적을 계산하고, 각 클래스별 AP의 평균인 mAP(mean Average Precision)를 평가척도로 활용한다.

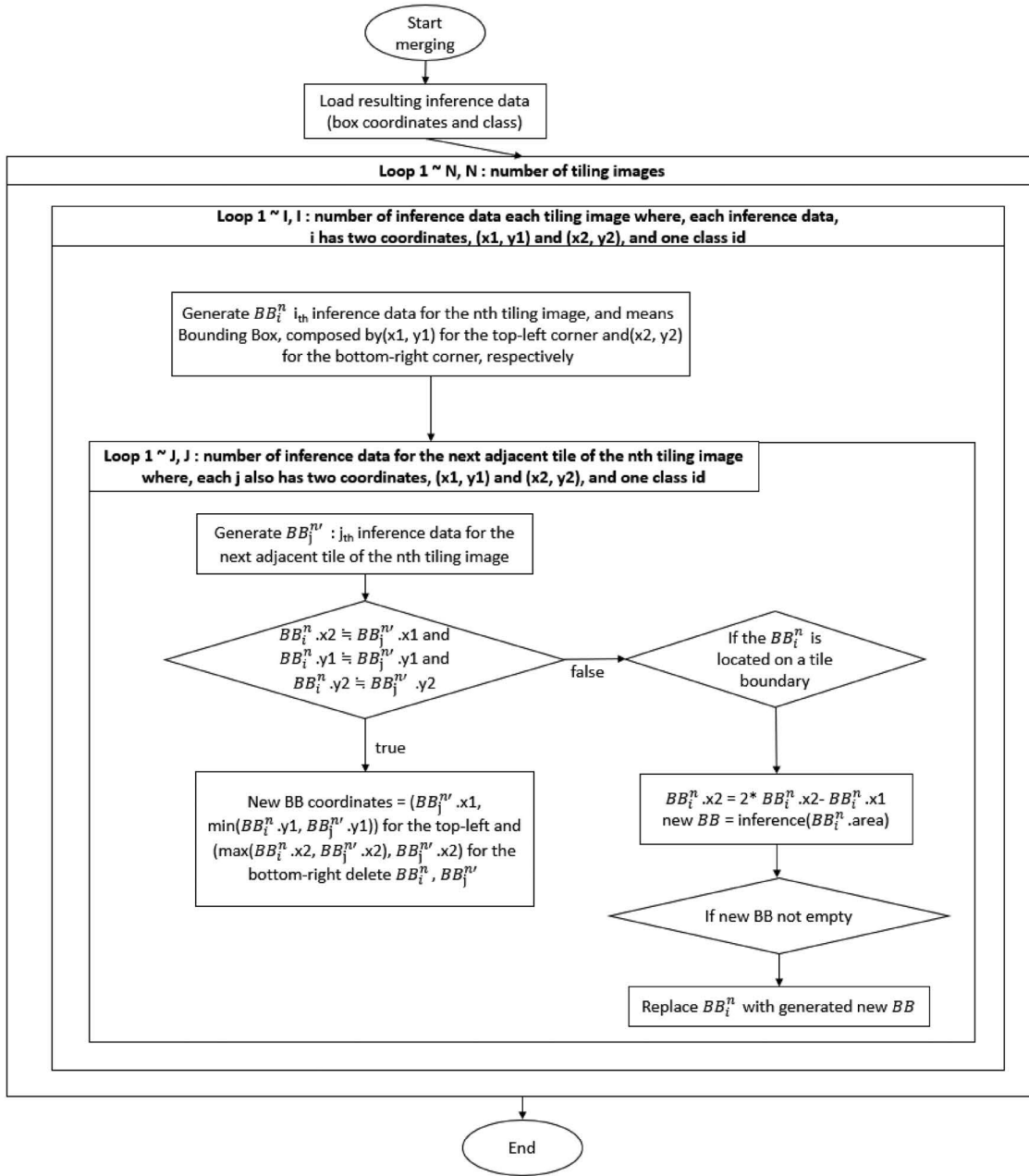


Fig. 7. Proposed bounding box merging algorithm.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

4. YOLOX를 이용한 말별 데이터 학습

YOLOX는 앵커 박스를 전혀 사용하지 않는 객체 검

출기이다. Table 3은 YOLO의 버전에 따른 성능과 특성을 나타낸다. COCO dataset에 대한 mAP의 경우, COCO가 YOLOv1 (Redmon et al., 2016) 이후에 나온 평가방식이라 v1을 제외했으며, 추론 시간에 대한 계산은 개발자의 github에 제시된 데이터이며 영상 1장에 대한 결과이다. YOLOv3부터 제시되어 v1, v2 (Redmon et al., 2017) 결과는 마찬가지로 제외하였다. YOLOv1~v4는 앵커 박스를 구성하는 크기 값과 가로세로 비율 값을 딥러닝 설

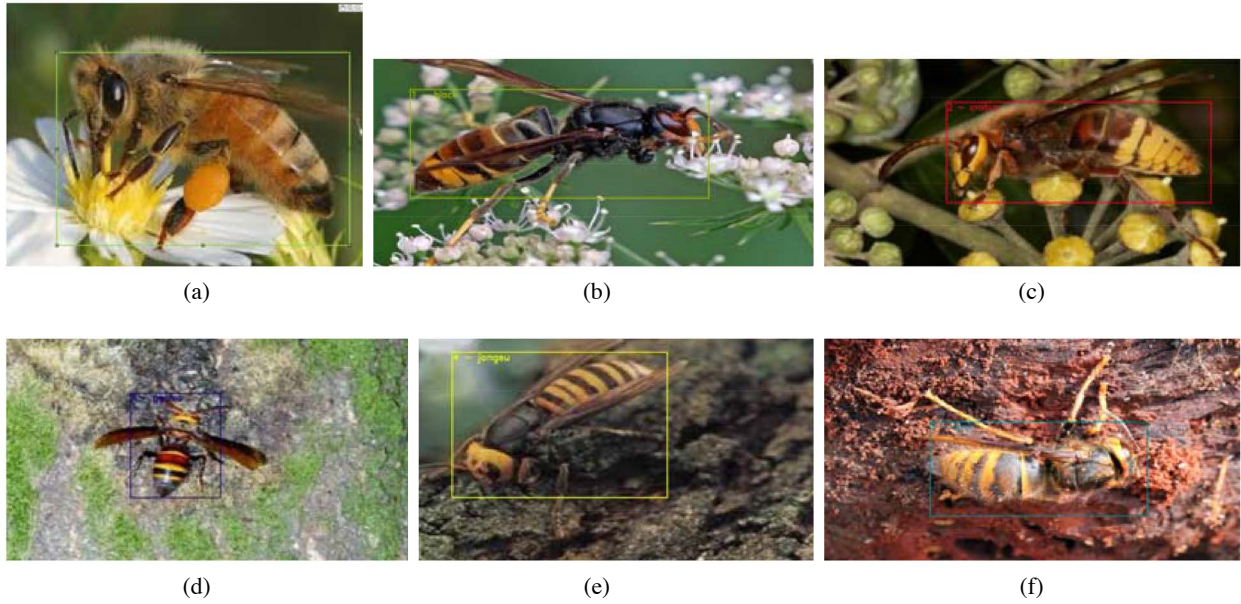


Fig. 8. Labeling for training data. (a) *Apis mellifera*, (b) *V. velutina nigrithorax*, (c) *V. crabro*, (d) *V. ducalis*, (e) *V. mandarinia*, (f) *V. similima*.



Fig. 9. Process to generate for test image set. (a) *Vespa* images, (b) Inference by YOLOX, (c) Crop the detected results, (d) Placing the cropped images on the boundary area of tiling.

게자가 직접 사전 정의하며 이에 따라 다양한 형태의 객체에 대한 모델의 유연성이 낮다. 그리고 학습 중에 발생한 모든 앵커 박스를 Ground truth에 대조해야 하므로 오버헤드가 발생하고 불필요한 negative 샘플을 과다 생성하는 부작용이 있다. 예를 들어 YOLOv4는 총 9개의 앵커 박스를 사전에 정의하므로 매번 9회 연산을 수행한다. 반면에 YOLOX를 포함한 앵커 프리 모델에서는 Fig. 10과 같이 그리드 셀의 왼쪽 모서리 상단인 중심점과 Ground truth 상자의 높이와 너비의 계산에 필요한 4개의 직선 길잇값을 직접 예측하도록 한다. 추가로 앵커 프리 모델은 작은 객체에 대해서 기존 앵커 기반 모델 대비 이점이 있다. 작은 객체를 대상으로 앵커 박스가 생성될 확률은 현저히 낮으며, 생성된 박스가 IOU (Intersection of Union)의 임계점을 넘지 못해 대부분 negative로 할당된다. 반면에 앵커 프리 모델은 예측된 중심점이 Ground truth 상자 내

Table 3. Comparison of YOLO-based detection networks.

YOLO-based networks	Release year	mAP for COCO dataset (%)	Inference time/image (ms)	Anchor
YOLO v1	2016	-	-	Anchor based
YOLO v2	2017	21.6	-	
YOLO v3	2018	33.0	21.9	
YOLO v4	2020	43.5	20.11	
YOLOX-s	2021	40.5	9.8	Anchor free
YOLOX-m		46.9	12.3	
YOLOX-l		49.7	14.5	
YOLOX-x		51.2	17.3	

부에 존재한다면 일단 positive로 할당되므로 유리하게 작동한다. 따라서 영상에서 대상 객체의 크기가 상대적으로 작은 형태인 말벌의 검출을 위해 처리 속도는 상대적으로

낮지만, 검출 성능이 우수한 YOLOX를 적용하였다.

딥러닝 학습에 사용된 하드웨어 환경은 Intel I9-9940X CPU, 64GB RAM, RTX 2080ti (x2) GPU이며 소프트웨어 환경은 Ubuntu 18.04 LTS, Python 3.8.12, CUDA 10.2, CUDNN 8.2.1, pytorch 1.10.1, torchvision 0.11.2, openCV 4.5.4이다. YOLOX는 신경망 네트워크의 깊이, 규모에 따라 모델을 세분화했다. 가벼운 순서대로 Nano, Tiny, s, m, l, x로 구성되어 있으며 본 연구에서는 가장 보편적인 model-s를 사용해 학습했다. 그 밖에 학습 환경으로 FP16 적용(Micikevicius *et al.*, 2017), 배치 사이즈 16, epoch 300, ImageNet (Krizhevsky *et al.*, 2012)을 사전 학습한 yolox_s.pth 전이 학습을 적용했다. Table 1의 7,505개의 데이터에서 test 900개를 제외한 6,605개의 영상과 레이블링 정보를 학습에 사용했다. 본 논문에서는 성

능 평가의 표준이 되는 PASCAL VOC (Everingham *et al.*, 2010)를 활용하였다. 이 방식은 구체적으로 예측된 결과와 ground truth의 IOU값이 0.5보다 크다면 Positive, 작다면 Negative로 분류하여 mAP를 계산한다. 모든 테스트 영상에 대해 mAP를 일괄 계산하고 시각화하고자 object detection metrics (Padilla *et al.*, 2021)를 사용했다.

결과 및 고찰

1. 타일링 방법에 따른 정량적인 말벌 인식 성능 비교

경계영역에서 말벌이 집중적으로 배치된 150장의 테스트 영상에 대해 mAP를 측정하였다. 첫 번째는 타일링

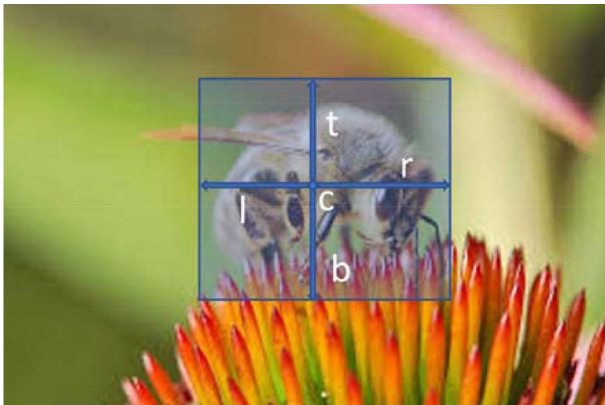


Fig. 10. Anchor-free model for the center point, c and 4 values with l, t, r, and b.

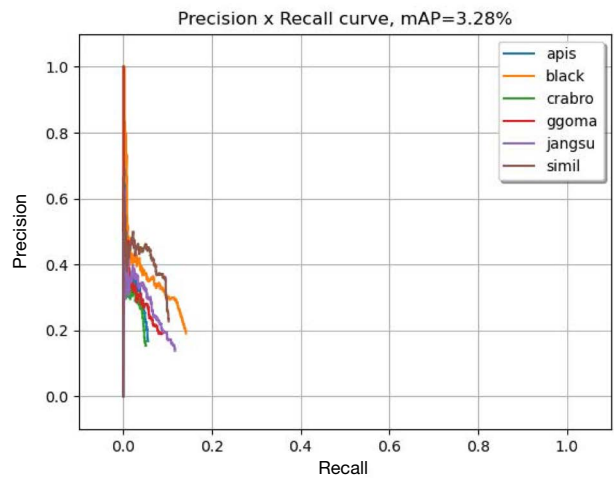


Fig. 12. mAP according to PR curve by Stephane's tiling.

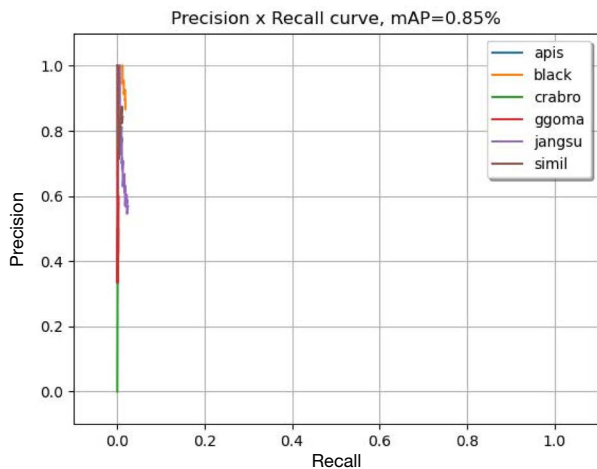


Fig. 11. mAP according to PR curve without tiling.

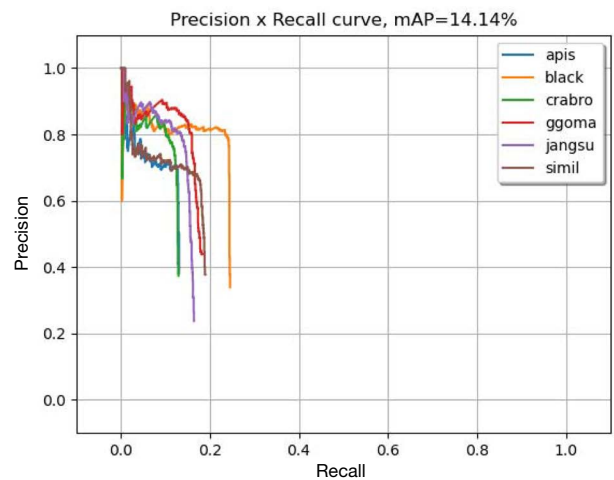


Fig. 13. mAP according to PR curve by proposed tiling.

Table 4. mAP result table for compared algorithms.

	AP according to classes (%)						mAP (%)
	Apis	Black	Crabro	Ggoma	Jangsu	Simil	
Without tiling	0.15	1.79	0.07	0.15	1.85	1.09	0.85
Stephane's tiling	2.09	5.31	1.65	2.70	3.47	4.45	3.28
Proposed tiling	9.93	20.64	10.89	15.65	13.65	14.07	14.14

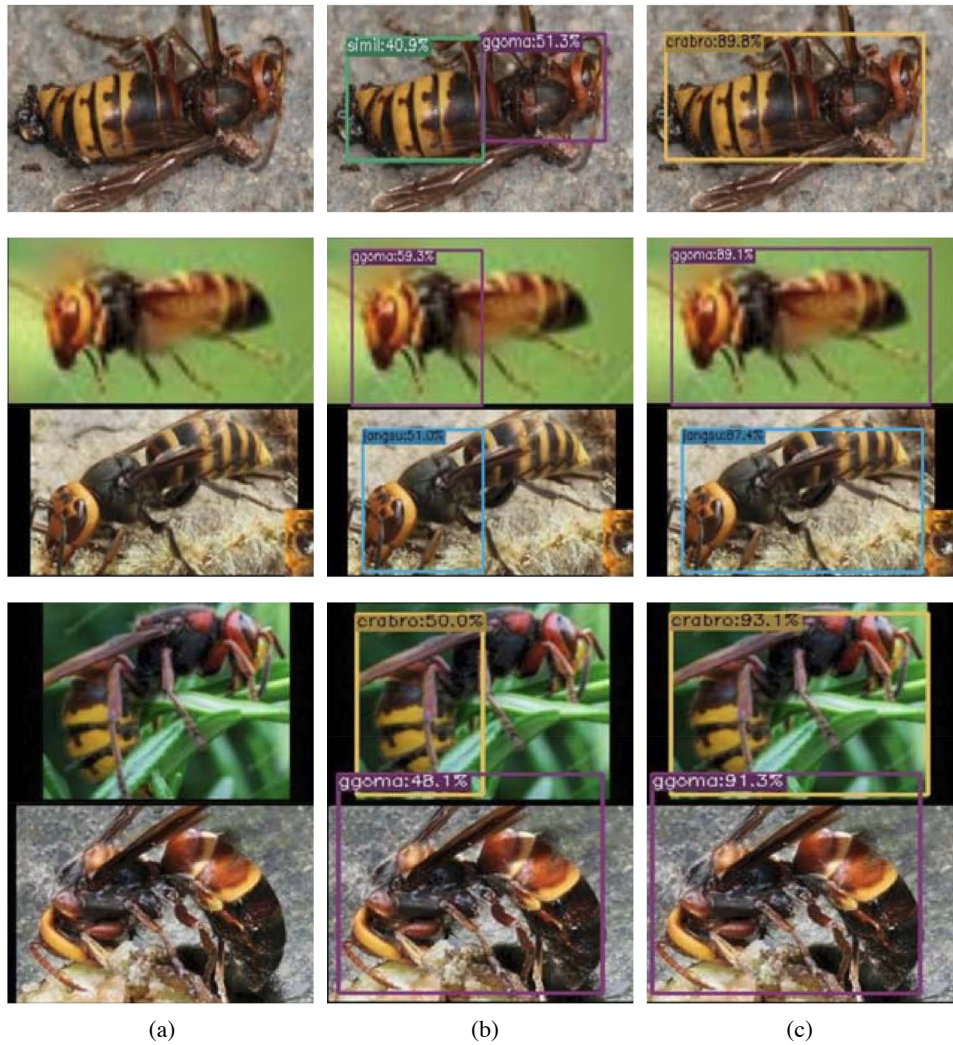


Fig. 14. Comparison for vespa detection. (a) Resultant images without tiling, (b) Resultant images by Stephane's tiling, (c) Resultant images by proposed tiling.

을 적용하지 않고 원 영상 그대로 YOLOX로 객체를 탐지한 후 mAP를 측정하였다. 두 번째는 Stephane의 타일링 알고리즘과 YOLOX를 이용하여 객체를 탐지하고 mAP를 측정하였다. 세 번째는 본 논문에서 제안한 개선된 타일링 방법을 적용하여 YOLOX로 객체를 탐지한 후 mAP

를 측정한 결과이다. 실행 시 사용한 파라미터 값은 conf 0.25, tsize 320이다. Conf 값은 검출 결과와 ground truth 사이의 IOU가 아무리 높아도 confidence가 기준치 미만이면 해당 결과는 사용하지 않는 옵션이다. Fig. 11~Fig. 13과 Table 4는 PR 커브에 따른 mAP의 측정 결과를 나타낸다.



Fig. 15. Examples of detection error, when the size difference between adjacent bounding boxes is large in proposed method.

결과, Proposed tiling, Stephane's tiling, Without tiling 순으로 mAP의 성능이 우수했다. 6가지의 개별 클래스에도 모두 동일한 순서로 결과가 확인되었다. 타일링을 적용하지 않았을 때 mAP가 지나치게 낮은 결과를 보이는 이유는 테스트 데이터의 환경에 있다. 타일링 테스트용 영상은 4K 해상도를 가지고 있으며 그 안에 150~200 px의 벌 영상이 연속으로 배치되어 있다. 심층 합성곱 신경망의 입력 크기를 맞추기 위해 입력 영상은 320×320 해상도의 정사각형으로 재조정되고 따라서 영상에 포함된 말벌은 10~20화소 정도로 작아지기 때문에 바운딩 박스의 생성이 어려워진다. 이뿐만 아니라, 작은 객체일수록 어긋난 1화소로 인한 인식률 저하가 크게 나타난다. 타일링을 적용하면 9개의 영역 분할을 진행하므로 영상의 크기 조정 후 말벌의 타일 영상에서 대상 객체의 크기는 상대적으로 크게 유지된다. 따라서 더 많은 바운딩 박스가 생성되고 mAP값이 높아진다.

2. 타일링 방법에 따른 테스트 결과 영상 비교

Fig. 14에서 (a)는 타일링 미적용 환경에서 테스트 데이터를 YOLOX로 인식한 결과이며, Fig. 14(b)는 Stephane의 타일링 방법을 사용하여 테스트 데이터를 인식한 결과이고 Fig. 14(c)는 본 논문이 제안한 방법으로 테스트 데이터를 인식한 결과의 영상의 일부이다. 앞서 언급한 정량적인 mAP결과와 마찬가지로 제안한 방법을 이용한 결과가 가장 우수한 성능을 보임을 영상으로도 확인할 수 있다. Fig. 14(a)와 같이 타일링을 사용하지 않은 검출에서 mAP가 매우 낮게 나온 점은 객체를 잘못 인식했기 때문이 아니라, 바운딩 박스가 아예 등장하지 않았기 때문임을 Fig. 14(a)로부터 확인할 수 있다. Fig. 14(b)의 맨 위 사진은 박스가 나타났지만 말벌에 대해 털보 말벌과 꼬마 말벌이라고 인식하여 객체를 추론하는 데 실패했으며, 각

각 40.9%, 51.3%로 낮은 확률을 보였지만 본 논문이 제시하는 알고리즘에서는 정상적으로 말벌로 인식했으며 그 확률도 89.8%를 얻었다. 또한 Fig. 14(b)의 가운데 3장 영상에서 Stephane의 타일링은 바운딩 박스가 누락되어 절반만 나타났지만, 본 논문에서 제시하는 알고리즘은 정상적으로 처리했음을 알 수 있다. 그러나 Fig. 15와 같이 제안한 방법에서도 인접하는 바운딩 박스가 서로 크게 어긋났을 때 영상에 포함된 객체의 위치와 종류에 대한 충분한 검출이 이루어지지 않음을 확인할 수 있다. 따라서 향후 연구에서는 이러한 오류에 대한 지속적인 개선 연구가 필요함을 확인하였다.

적 요

양봉 산업의 주요 피해 요인인 말벌의 효과적인 방제를 위해서는 말벌의 출현을 확인하기 위한 실시간 모니터링 시스템의 도입이 필요하다. 현재 영상인식 분야에서는 합성곱 신경망이 많이 활용되고 있으나 합성곱 신경망을 활용한 객체 인식을 위해서는 충분한 객체의 크기가 보장되어야 한다. 40 cm 거리에서 4K 카메라로 가로 세로 각각 30~40 cm의 영역을 촬영할 경우 전체 영상에서 말벌의 크기는 2~3% 내외로 화소 수로는 가로세로 대략 100~120화소 이내이다. 따라서 인식이 매우 어려워 영상을 분할하여 최소 입력 크기가 보장될 수 있도록 하는 타일링 방법이 필요하다. 그러나, 기존의 타일링 방식은 분할된 경계선 영역에서 객체가 다르거나, 절반만 검출되었을 때 대상을 잘 분류하지 못하는 한계점이 있는 것으로 확인되었다. 따라서 본 연구에서는 기존 타일링 방법에서 객체가 다르거나 절반만 객체로 검출될 때 추가적인 분류를 포함하는 형태로 알고리즘을 개선하여 인식 성능을 향상하는 방법을 제시하였다. 제안한 방법의 성능평가를 위

해 5종의 말벌과 1종의 양봉꿀벌 영상으로 구성된 7,505 개의 데이터와 YOLOX를 이용한 말벌 검출 시스템을 개발하였다. 성능평가를 위해 4K 해상도를 가진 배경에 다수의 말벌 영상을 타일 영상의 경계영역에 겹치지 않게 합성한 시험 영상을 제작하여 경계영역에 존재하는 말벌 객체의 인식 성능을 평가하였다. 평가도구로는 객체 탐지 척도(Padilla *et al.*, 2021)를 활용하였으며 실험 결과 제안한 방법이 PASCAL VOC mAP 기준 14.14%로 정확도의 측면에서 비교 알고리즘보다 우수한 결과를 보임을 확인하였다.

감사의 글

본 논문은 농촌진흥청 공동연구사업(과제번호: PJ0147 6103)으로 수행된 결과이며 농촌진흥청의 지원에 깊은 감사를 드립니다.

인용 문헌

- Bai, Y., Y. Zhang, M. Ding and B. Ghanem. 2018. Finding tiny faces in the wild with generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition 21-30.
- Bochkovskiy, A., C. Y. Wang and H. Y. M. Liao. 2020. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.
- Choi, M. B., J. K. Kim and J. W. Lee. 2013. Checklist and Distribution of Korean Vespidae Revisited. Korean J. Appl. Entomol. 52(2): 85-89.
- Everingham, M., L. Van Gool, C. K. Williams, J. Winn and A. Zisserman. 2010. The pascal visual object classes (voc) challenge. IJCV 88(2): 303-338.
- Ge, Z., S. Liu, F. Wang, Z. Li and J. Sun. 2021. Yolox: Exceeding YOLO Series in 2021. arXiv preprint arXiv:2107.08430.
- Jeong, Y. J., Y. H. Lee, I. Ansari and C. H. Lee. 2020. Real Time Hornet Classification System Based on Deep Learning. Journal of IKEEE. 24(4): 1141-1147.
- Jung, C. E. 2012a. Spatial Expansion of an Invasive Hornet, *Vespa velutina nigrithorax* Buysson (Hymenoptera: Vespidae) in Korea. Korean J. Apic. 27(2): 87-93.
- Jung, C. E. 2012b. Initial stage risk assessment of an invasive hornet, *Vespa velutina nigrithorax* Buysson (Hymenoptera: Vespidae) in Korea. Korean J. Apic. 27: 95-104.
- Jung, C. E., D. W. Kim, H. S. Lee and H. Baek. 2009. Some Biological Characteristics of a New Honeybee Pest, *Vespa velutina nigrithorax* Buysson, 1905 (Hymenoptera: Vespidae) in Korea. Korean J. Apic. 24(1): 61-65.
- Jung, C. E., M. S. Kang, D. W. Kim and H. S. Lee. 2007. Vespidae wasps (Hymenoptera) occurring around apiaries in Andong, Korea: II. Trap catches and seasonal dynamics. Korean J. Apic. 22: 63-70.
- Kim, B. S., S. M. Jeong, G. E. Kim and C. E. Jung. 2017. Early alert system of Vespa attack to honeybee hive: prototype design and testing in the laboratory condition. Korean J. Apic. 32(3): 191-198.
- Kisantal, M., Z. Wojna, J. Murawski, J. Naruniec and K. Cho. 2019. Augmentation for small object detection. arXiv preprint arXiv:1902.07296.
- Krizhevsky, A., I. Sutskever and G. E. Hinton. 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.
- Lee, C. H., Y. J. Jeong, T. H. Kim, J. H. Park, S. B. Park and C. Jeong. 2019. Performance Comparison of Deep Convolutional Neural Networks for Vespa Image Recognition. Korean J. Apic. 34(3): 207-215.
- Lin, T. Y., P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie. 2017. Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition 2117-2125.
- Micikevicius, P., S. Narang, J. Alben, G. Diamos, E. Elsen, D. Garcia, B. Ginsburg, M. Houston, O. Kuchaiev, G. Venkatesh and H. Wu. 2017. Mixed precision training. arXiv preprint arXiv:1710.03740.
- Padilla, R., W. L. Passos, T. L. Dias, S. L. Netto and E. A. da Silva. 2021. A comparative analysis of object detection metrics with a companion open-source toolkit. Electronics 10(3): 279.
- Redmon, J. and A. Farhadi. 2017. YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition 7263-7271.
- Redmon, J. and A. Farhadi. 2018. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.
- Redmon, J., S. Divvala, R. Girshick and A. Farhadi. 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition 779-788.
- Stephane, C. <https://github.com/stephanecharette/DarkHelp>