

Identification and Area Calculation of the Egg Mass of Spodoptera Frugiperda Based on Faster-RCNN

Zhishang Liang

School of Electronics and Information Engineering, GuangDong Ocean University

Yinqiao Peng (✉ yinqiaopeng@163.com)

School of Electronics and Information Engineering, GuangDong Ocean University

Muhammad Sameer Sheikh

School of Electronics and Information Engineering, GuangDong Ocean University

Zhengwei Wu

College of Coastal Agricultural Sciences, GuangDong Ocean University

Ji Wang

School of Electronics and Information Engineering, GuangDong Ocean University

Research Article

Keywords: Spodoptera frugiperda, Identification of egg, Estimation of area, Faster-RCNN, Maximum Between-Class Variance method

Posted Date: December 6th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1109249/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

32 gray villous protective layer ^{5,6}. The main prevention and control measures include
33 monitoring technology, agricultural control, physicochemical control, biological
34 control and chemical control. Accurate monitoring based on local insect situation and
35 migration can effectively control the harm of *Spodoptera frugiperda* ⁷. The number of
36 egg granules is an important criterion to evaluate and predict the damage degree of
37 insect pests. The traditional method of counting egg granules by hand has low efficiency,
38 high labor intensity and low accuracy. Friedland et al. ⁸ used the image processing
39 software Optimas to calculate the fish eggs without overlap, and the error rate in the
40 experimental sample was 1%, but this method needed to set the threshold of image
41 segmentation artificially. Duan et al. ⁹ used threshold segmentation, morphological
42 operation and watershed segmentation to segment egg images and count the eggs of
43 pelagic fish in the MATLAB programming environment. Yang et al. ¹⁰ used background
44 estimation, morphological image processing, background subtraction, stretching, image
45 thresholding, gray-level transformation, labeling and counting methods to
46 automatically detect and count mantispid eggs. Zhang et al. ¹¹ proposed a method of
47 shrimp egg counting based on convolutional neural network, built a shrimp egg
48 counting device based on computer vision, and developed a computer operating
49 program to count shrimp eggs in images. The average counting accuracy of this method
50 is 99.2%. LI Zi-Yuan et al. ¹² proposed a simple method to estimate the number of eggs
51 in the egg mass of *spodoptera frugiperda*, and established an exponential function
52 model of eggs number and egg mass area, which provided a reference for the rapid
53 estimation of the number of eggs. There is a positive correlation between the number
54 of eggs and the area of egg mass. It is an effective method to estimate and count the
55 number of eggs by measuring the area, but it is inefficient to measure and calculate the
56 area of egg mass by hand. Based on the function model of egg number and egg mass
57 area, this paper uses deep learning method and image processing technology to obtain
58 the egg area of *spodoptera frugiperda* and quickly estimate the egg number.

59 Deep learning is a new research direction in the field of machine learning ¹³. In
60 recent years, it has become an effective method of feature representation and achieved
61 significant breakthroughs in the field of target detection. Target detection is to extract
62 the target from the image and then determine target classification ¹⁴. Target detection
63 algorithms are mainly divided into target detection based on candidate regions, such as
64 RCNN, Fast-RCNN and Faster-RCNN ¹⁵, and target detection based on regression
65 method, such as YOLO ¹⁶. Fu et al. ¹⁷ adopted the improved YOLOV3-Tiny model to

66 realize the rapid detection of kiwifruit on small hardware devices, while ensuring the
67 real-time detection. Li et al. ¹⁸ respectively adopted Faster-RCNN, YOLOV3 and SSD
68 to detect the target of agricultural greenhouses in high spatial resolution images. Wu et
69 al. ¹⁹ adopted the YOLOV3 deep learning algorithm and the classification method of
70 lame and non-lame cows based on the relative step size vector, and realized the
71 intelligent detection of lame cows. Yarak et al. ²⁰ used Faster-RCNN for automatic
72 detection and health classification of oil palm, and compared the Faster-RCNN model
73 using two deep full convolutional networks, Resnet50 and VGG16, respectively. Jin et
74 al. ²¹ used Fast-RCNN to detect the cross section of wheat stem in the field and realized
75 the detection of wheat spike density. Zhou et al. ²² adopted the improved Faster-RCNN
76 algorithm to realize strawberry target detection in ground RGB images, and the average
77 accuracy reached more than 86%. Li et al. ²³ proposed an improved lightweight target
78 detection method based on YOLOv3, which realized the identification of typical
79 obstacles in the orchard, such as people, cement columns, telephone poles, etc. The
80 above studies provide a reference for the application of target detection network in the
81 field of agriculture. However, there are few studies on the egg mass of agricultural pests
82 by deep learning algorithm.

83 In this paper, the Faster-RCNN target detection network was used to identify the
84 eggs of *spodoptera frugiperda* in the image, and the deep full convolutional network
85 VGG16 in Faster-RCNN was replaced with Resnet50 to optimize the model. Then the
86 identified egg was binarized by the Maximum Between-Class Variance method (Otsu)
87 to get the exact size and shape of the egg, and finally the area of the egg was calculated.

88 **2. Materials and Methods**

89 The data set in this paper was obtained by photographing the egg mass of the
90 captive reared *spodoptera frugiperda* in the laboratory. Under the condition of sufficient
91 and uniform light, the collected eggs were placed on a tray, and the eggs were
92 photographed with a camera. There are 844 images of egg mass in the dataset, of which
93 458 have a resolution of 5632*4224, 38 have a resolution of 3024*4032, 348 have a
94 resolution of 3456*4608. The sample image of egg mass collected is shown in Fig. 1.

95 In order to enhance the anti-interference ability and generalization ability of the
96 model, part of images in the data set were randomly selected for data augmentation
97 before the training data set, including random rotation, flipping, stretching, etc.
98 Through the data augmentation method, the images of the data set were increased from
99 844 to 949. The comparison before and after the data augmentation is shown in Fig. 2.

100 Before training, a rectangular frame was manually labeled on the area where the
101 egg mass was in each image with the Labeling tool to obtain the real frame. The egg
102 mass was labeled as EGG, and an image example after Labeling was shown in Fig. 3.

103 Faster-RCNN¹⁵ is the third generation of RCNN series algorithm proposed by
104 Ren Shaoqing et al in 2016. The structure of Faster-RCNN shown in Fig. 4 includes
105 deep full convolutional network, Region Proposal Networks (RPN), ROI Pooling
106 module, target classification and positioning module²⁴⁻²⁶.

107 The commonly used deep full convolutional network for Faster-RCNN is VGG16,
108 whose structure is shown in Fig. 5. In the VGG16 network structure^{27,28}, there are 13
109 convolution layers, all of which adopt 3×3 convolution kernel with step size of 1, which
110 can increase the depth of the network and reduce the number of total parameters at the
111 same time. A nonlinear activation layer with ReLU is usually carried out after each
112 convolution layer. The pooling layer is used for subsampling and adopts the pooling
113 core of 2×2, the step size of 2, and the filling method of SAME. After the image passes
114 through the deep full convolutional network, the feature map is generated.

115 The algorithm structure of Region Proposal Networks (RPN) is shown in Fig. 6.
116 Region Proposal Networks (RPN) generates anchor boxes of different sizes and
117 proportions on the pixels of feature map generated by the deep full convolutional
118 network, and then uses these anchor boxes and ground-truth boxes in the picture to
119 calculate Intersection over Union (IoU). Non-maximum suppression is used to
120 eliminate overlapped anchor boxes. For the rest of the anchor boxes, the classification
121 probability is calculated using the SoftMax function, and the position and size are
122 adjusted using the regression method to get the proposed boxes.

123 The ROI pooling module combines the proposed box of different sizes generated
124 by RPN with the feature map generated by the deep full convolutional network. The
125 mapping is formed on the feature map according to the size and position information
126 of the proposed box, and then the maximum pooling operation is carried out to generate
127 the feature vector of the same size, which contains the size, position and classification
128 information of the proposed box. Finally, the full connection network is used to
129 complete the final classification and positioning.

130 The loss function of Faster-RCNN¹⁵ is defined as

$$131 \quad L(\{P_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(P_i, P_i^*) + \lambda \frac{1}{N_{reg}} \sum_i P_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

132 Here, i is the index value in a batch, and P_i is the prediction probability of the

133 ith anchor box. The value of P_i^* is 1 if the ground-truth tag in the real bounding box
 134 represents a positive example, otherwise its value is 0. t_i is the vector of the prediction
 135 bounding box, and t_i^* is the vector of the positive example ground-truth bounding box.
 136 L_{cls} is the classification loss, L_{reg} is the regression loss. N_{cls} and N_{reg} are
 137 respectively the number of images and the number of anchor frames in each small batch.

138 In the deep learning task, due to the complexities of the shape, color and other
 139 features of the target in the image, a deeper convolutional neural network is needed. As
 140 the depth of the network increases, the accuracy reaches saturation and then rapidly
 141 degrades. The deeper the network is, the greater the training error and testing error will
 142 be, and the gradient will disappear or explode. In order to solve this problem, He
 143 Kaiming et al.²⁹ introduced residual structure and proposed deep residual network
 144 ResNet. Deep residual network has been widely used due to its remarkable parameter
 145 optimization capability³⁰⁻³¹. The deep residual network is composed of residual units,
 146 which are divided into 2-layer residual units and 3-layer residual units, as shown in Fig.
 147 7.

148 Resnet has a variety of network structures, including 18 layers, 34 layers, 50
 149 layers, 101 layers and 152 layers. Resnet18 and Resnet34 use 2-layer residual units,
 150 and Resnet50, Resnet101 and Resnet152 use 3-layer residual units. The structure of the
 151 RESNET-N network is shown in Table 1. Considering the accuracy of defect detection
 152 and actual computing ability, this paper uses Resnet50 to replace the deep full
 153 convolutional network VGG16 in Faster-RCNN algorithm to optimize the model.

154 After using the target detection algorithm to get the position information of the
 155 egg mass in the picture, it is necessary to further calculate the proportion of the egg
 156 mass in the whole picture to calculate the more accurate area of the egg mass. The
 157 output box in the image is intercepted, and then the intercepted image is converted into
 158 a binary image. Image binarization is to convert digital images into black and white
 159 images, which is a simple thresholding process. Pixels whose gray level is lower than
 160 a given threshold are divided into one class, and the remaining pixels are divided into
 161 another class. The segmentation algorithm can be expressed as

$$162 \quad g(x, y) =$$

$$163 \quad \begin{cases} 255, & f(x, y) < T \\ 0, & f(x, y) > T \end{cases} \quad (2)$$

164 Here, $f(x,y)$ is the grayscale value of coordinate points whose coordinates are (x,y)
 165 in the gray image, T is the segmentation threshold, and $g(x,y)$ is the grayscale value of

166 coordinate points whose coordinates are (x,y) in the image after binarization. Selecting
 167 the size of threshold T is the key to image segmentation.

168 Many image binarization methods have been proposed, among which the most
 169 classic is the Maximum Between-Class Variance method (Otsu) ³². According to the
 170 gray characteristics of the image, the image is divided into foreground and background.
 171 When the between-class variance between foreground and background is the largest, it
 172 indicates that the difference between foreground and background is large, and the
 173 threshold is the best. For the image to be processed, the probability distribution of
 174 grayscale is

$$175 \quad \begin{cases} P_i = \frac{n_i}{N} \\ \sum_{i=0}^{L-1} P_i = \frac{n_0+n_1+\dots+n_{L-1}}{N} = 1 \end{cases} \quad (3)$$

176 Here, n_i is the number of pixels with gray value i in the image, N is the total
 177 number of pixels in the image, L is the gray level of the image, and P_i is the probability
 178 of pixels with gray value i .

180 A threshold value t is selected to divide the gray level in the image into two parts,
 181 $C_0=(0,1,2, t)$ and $C_1=(t+1,t+2, L-1)$. The probability of both can be expressed as
 182 ω_0

$$183 \quad \omega_0 = \sum_{i=0}^t P_i$$

$$184 \quad \omega_0 = \omega(t) \quad (4)$$

$$185 \quad \omega_1 = \sum_{i=t+1}^{L-1} P_i = 1 - \omega(t) \quad (5)$$

187 The mean of the probabilities can be expressed as

$$188 \quad \mu_0 = \sum_{i=0}^t \frac{iP_i}{\omega_0} = \frac{\mu(t)}{\omega(t)} \quad (6)$$

$$189 \quad \mu_1 = \sum_{i=t+1}^{L-1} \frac{iP_i}{\omega_1} = \frac{\mu_T(t)-\mu(t)}{1-\omega(t)} \quad (7)$$

193 The variance can be expressed as:

194 σ_0^2

195
$$= \sum_{i=0}^t \frac{(i - \mu_0)^2 P_i}{\omega_0}$$
 (8)

196 $\sigma_1^2 =$

197
$$\sum_{i=t+1}^{L-1} \frac{(i - \mu_1)^2 P_i}{\omega_1}$$

198 (9)

199 The inter-class variance can be expressed as:

200
$$\sigma_A^2 = \omega_0(\mu_1 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 = \omega_0\omega_1(\mu_1 - \mu_0)^2$$
 (10)

201 Each number between 0 and 255 is selected as the selected threshold t to calculate
 202 the between-class variances respectively. The t value with the largest between-class
 203 variance is selected as the threshold T of image segmentation, and the image is
 204 transformed into a binary image. After the intercepted image is converted into a binary
 205 image, the position, size and shape of the egg mass are determined in the original image
 206 according to the size and position of the output box of Faster-RCNN and the converted
 207 binary image. In the original image, the part of the egg mass is represented by white
 208 pixels, while the rest is represented by black pixels. The area of the egg mass can be
 209 calculated using the proportion of the white part in the original image. The egg mass
 210 area is calculated as follows:

211 *white_percnet*

212
$$= \frac{white_num}{all_num}$$
 (11)

213 $S = white_percent \times$

214 *truth_area* (12)

215 Here, *white_num* is the number of white pixels in the picture, *all_num* is the
 216 number of pixels in the picture, *white_percent* is the percentage of white pixels in the
 217 picture, *truth_area* is the real area of the picture, and S is the calculated egg area.

218 In order to verify the accuracy of the calculation results, the length and width of
 219 the egg mass were measured manually and the real area of the egg mass was calculated.
 220 The egg mass of *Spodoptera frugiperda* is irregular in shape. In order to make the
 221 measured data closer to the real value and reduce the error in the measurement process,
 222 the measurement is carried out at three different points when measuring the length and
 223 width, and each point is repeated three times. Then average the data. The egg mass area
 224 is calculated according to the approximate elliptical area formula [12], and the egg mass

225 area is expressed as

$$226 \quad X = L \times W \times \frac{\pi}{4} \quad (13)$$

227 Here, L is the length of the egg mass, W is the width of the egg mass, and X is the
228 area of the egg mass.

229 In order to obtain the true area of the original image, the square drawing paper is
230 used as the background of the image. Place the egg mass to calculate the area on the
231 square drawing paper. For ease of calculation, the resolution of the camera image is set
232 to 1:1 for shooting, and the resolution of the image obtained is 4224*4224. As shown
233 in Fig. 8, the size of each small square in the figure is 1mm * 1mm, and the real area of
234 the whole picture can be obtained by counting the number of squares.

235 In summary, after the output frame is obtained by using Faster-RCNN, the image
236 in the output frame is divided, and then formula 11, 12 and 13 are used to calculate the
237 area of egg blocks.

238 **3. Results**

239 In the training phase of Faster-RCNN, 90 percent of the images are used for
240 training and the remaining 10 percent for validation. The initial learning rate was
241 selected as 0.005, the attenuation coefficient was selected as 0.33, the batch size was
242 selected as 2, and the epoch was selected as 100. The experimental computer was
243 configured with NVIDIA GeForce RTX 2060 GPU, and the model was built using
244 Python based on PyTorch deep learning framework.

245 Common model evaluation indexes in target detection models include accuracy,
246 precision, recall, etc. [18,24]. Accuracy is the proportion of the number of correct
247 predictions to the number of observations. Precision is the proportion of correctly
248 predicted samples in all positive samples. The recall rate is the proportion of correct
249 predictions in all positive samples. Since recall rate and precision rate were
250 contradictory, the model was evaluated using AP of average accuracy in the same
251 category and mAP of average accuracy in different categories. By lowering the
252 classification threshold, the recall rate and corresponding precision under each
253 threshold were calculated, the precision-recall curve was drawn, and the area enclosed
254 by the curve and the coordinate axis was calculated. The larger the area, the better the
255 prediction effect of the model would be. Since there is only one classification type in
256 the experiment in this paper, the value of AP and the value of mAP are equal.

257 In this paper, the performance of the model is evaluated comprehensively with

258 multiple indexes. In order to evaluate the performance of the improved model, the
259 improved model of deep full convolutional network replacing VGG16 with Resnet50
260 was compared with the original model in the training and testing process. In the training
261 process, the loss change process of training and validation is shown in Fig. 9. With the
262 increase of the number of training epochs, the errors of all models decreased
263 continuously, and the decline rate of the improved model was faster. After 100 epochs
264 of training, the minimum losses of the original model in the training set and the test set
265 were 0.640 and 0.656, respectively, while the minimum losses of the improved model
266 in the training set and the test set were 0.571 and 0.650, respectively. The mAP of
267 different models are shown in Table 2. The mAP of the improved model are all higher
268 than those of the original model. The accuracy of the improved model and the original
269 model were 80.82% and 80.64%, respectively. The precision-recall curves of different
270 models are shown in Fig. 10. The area of the improved model is larger than that of the
271 original model. An example of improved model detection results is shown in Fig. 11.
272 The comparison results of the above data indicate that the performance of the Faster-
273 RCNN model using RESNET50 as a deep full convolutional network is superior to the
274 original Faster-RCNN model.

275 After the egg mass quality and position in the image are obtained by the target
276 detection method, the image in the output box is intercepted, as shown in Fig. 12. The
277 intercepted image is converted into gray image, and then it is converted into binary
278 image by the Maximum Between-Class Variance method, as shown in Fig. 13. After
279 the part of egg mass in the original image is represented by white pixels and the
280 remaining part is represented by black pixels, the binary image of the original image is
281 obtained, as shown in Fig. 14. The egg mass area calculated by equations 11 and 12 and
282 the actual egg mass area calculated by equation 13 are shown in Table 3. As can be seen
283 from Table 3, the average relative error in 16 test samples is -0.02032, and the minimum
284 value is -0.00047, which can meet the requirements of egg mass area measurement.

285 **4. Conclusion**

286 In this paper, 844 images of egg mass of *spodoptera frugiperda* were collected,
287 and the data were augmented random flipping, rotating, stretching and other methods.
288 Resnet50 was used to replace VGG16 to optimize Faster-RCNN to detect egg masses
289 in the image. The Maximum Between-Class Variance method is used to convert the
290 image in the output box into a binary image, and then the area of egg quality is
291 calculated. The average relative error of egg mass area was -0.02032, and the minimum

292 value was -0.00047. Compared with the traditional manual method, the optimized
293 Faster-RCNN and the Maximum Between-Class Variance method were used to
294 calculate the egg mass area and estimate the egg number, which greatly improved the
295 efficiency and precision of egg mass statistics, and provided a new idea for the
296 evaluation of the severity of agricultural insect pests. In the industry, the detection of
297 targets requires real time. However, the processing time of the method in this paper is
298 about 3 seconds. Therefore, in the following research, we will look for a more rapid
299 detection method to improve its practicability.

300 **Acknowledgments**

301 The work was supported by Science and Technology Project of Guangdong
302 Province (No.2016A010101028) and 2019 Guangdong Education Department's
303 "artificial intelligence" key field special project (2019KZDZX1046).

304

305 **References**

- 306 1. He, Li-mei, et al. Larval diet affects development and reproduction of East Asian
307 strain of the fall armyworm, *Spodoptera frugiperda*. *Journal of Integrative*
308 *Agriculture*.**20**,736-744. [https://doi.org/10.1093/ee/nvaa172\(2021\)](https://doi.org/10.1093/ee/nvaa172(2021)).
- 309 2. Wang, Li-Kui, et al. Correction to: A Preliminary Toxicology Study on Eco-friendly
310 Control Target of *Spodoptera frugiperda*. *Bulletin of environmental contamination*
311 *and toxicology*.**106**,559. [https://doi.org/10.1007/s00128-021-03140-8\(2021\)](https://doi.org/10.1007/s00128-021-03140-8(2021)).
- 312 3. He, Limei, et al. Photoregime Affects Development, Reproduction, and Flight
313 Performance of the Invasive Fall Armyworm (Lepidoptera: Noctuidae) in
314 China. *Environmental Entomology*.**50**,367-381.
315 [https://doi.org/10.1093/ee/nvaa172\(2021\)](https://doi.org/10.1093/ee/nvaa172(2021)).
- 316 4. García-Gómez, Griselda, et al. Maize mycorrhizas increase the susceptibility of the
317 foliar insect herbivore *Spodoptera frugiperda* to its homologous
318 nucleopolyhedrovirus. *Pest Management Science*. <https://doi.org/10.1002/ps.6511>
319 (2021).
- 320 5. Boiça, Arlindo Leal, et al. *Spodoptera frugiperda* (JE Smith)(Lepidoptera:
321 Noctuidae) in cotton: vertical distribution of egg masses, effects of adult density
322 and plant age on oviposition behavior. *Arquivos do Instituto Biológico*. **80**,424-429.
323 [https://doi.org/10.1590/S1808-16572013000400008\(2013\)](https://doi.org/10.1590/S1808-16572013000400008(2013)).
- 324 6. Liao, Yong-Lin, et al. First report of *Telenomus remus* parasitizing *Spodoptera*
325 *frugiperda* and its field parasitism in southern China. *Journal of Hymenoptera*
326 *Research*.**79**,95. [https://jhr.pensoft.net/article/39136\(2019\)](https://jhr.pensoft.net/article/39136(2019)).
- 327 7. Yan, Z. H. O. U., et al. Spread of invasive migratory pest *Spodoptera frugiperda* and
328 management practices throughout China. *Journal of Integrative*
329 *Agriculture*.**20**,637-645.[https://doi.org/10.1016/S2095-3119\(21\)63621-3\(2021\)](https://doi.org/10.1016/S2095-3119(21)63621-3(2021)).
- 330 8. Friedland, K. D., et al. Automated egg counting and sizing from scanned images:
331 rapid sample processing and large data volumes for fecundity estimates. *Journal of*
332 *Sea Research*.**54**,307-316. [https://doi.org/10.1016/j.seares.2005.06.002\(2005\)](https://doi.org/10.1016/j.seares.2005.06.002(2005)).
- 333 9. Duan, Yane, et al. An automatic counting system for transparent pelagic fish eggs
334 based on computer vision. *Aquacultural engineering*.**67**,8-13.

- 335 <https://doi.org/10.1016/j.aquaeng.2015.05.001>(2015).
- 336 10. Yang, Pei-Ying, et al. Automatic mantispid egg detection and counting using image
337 nature. *International Journal of Modern Physics B.* **34**,2040138.
338 <https://doi.org/10.1142/S0217979220401384>(2020).
- 339 11. Zhang, Junjie, et al. Shrimp Egg Counting with Fully Convolutional Regression
340 Network and Generative Adversarial Network. *Aquacultural Engineering.*102175.
341 <https://doi.org/10.1016/j.aquaeng.2021.102175>(2021).
- 342 12. Li, Ziyuan, et al. A method for estimating the number of Spodoptera frugiper
343 (smith) eggs per egg mass. Transactions of the Chinese Society of Journal of
344 Environmental Entomology. **42**,1322-1329.doi:CNKI:SUN:KCTD.0.2020-06-
345 005 (2020).
- 346 13. Ning, Yishuang, et al. A review of deep learning based speech synthesis. *Applied*
347 *Sciences.* **9**,4050. <https://doi.org/10.3390/app9194050>(2019).
- 348 14. Lian, Jing, et al. Small object detection in traffic scenes based on attention feature
349 fusion. *Sensors.***21**,3031.<https://doi.org/10.3390/s21093031>(2021).
- 350 15. Ren, Shaoqing, et al. Faster r-cnn: Towards real-time object detection with region
351 proposal networks. *Advances in neural information processing systems.***28**,91-99.
352 <https://ieeexplore.ieee.org/document/7485869>(2015).
- 353 16. Redmon, Joseph, et al. You only look once: Unified, real-time object
354 detection. *Proceedings of the IEEE conference on computer vision and pattern*
355 *recognition.* (2016).
- 356 17. Fu, Longsheng, et al. Fast and accurate detection of kiwifruit in orchard using
357 improved YOLOv3-tiny model. *Precision Agriculture.***22**,754-776.
358 <https://doi.org/10.1007/s11119-020-09754-y>(2021).
- 359 18. Li, Min, et al. Agricultural greenhouses detection in high-resolution satellite images
360 based on convolutional neural networks: Comparison of faster r-cnn, yolo v3 and
361 ssd. *Sensors.* **20**,4938.<https://doi.org/10.3390/s20174938>(2020).
- 362 19. Wu, Dihua, et al. Lameness detection of dairy cows based on the YOLOv3 deep
363 learning algorithm and a relative step size characteristic vector. *Biosystems*
364 *Engineering.* **189**,150-163. <https://doi.org/10.1016/j.biosystemseng.2019.11.017>

- 365 (2020).
- 366 20. Yarak, Kanitta, et al. Oil Palm Tree Detection and Health Classification on High-
367 Resolution Imagery Using Deep Learning. *Agriculture*. **11**,183.
368 [https://doi.org/10.3390/agriculture11020183\(2021\)](https://doi.org/10.3390/agriculture11020183(2021)).
- 369 21. Jin, Xiuliang, et al. High-throughput measurements of stem characteristics to
370 estimate ear density and above-ground biomass. *Plant Phenomics*.2019.
371 [https://doi.org/10.34133/2019/4820305\(2019\)](https://doi.org/10.34133/2019/4820305(2019)).
- 372 22. Zhou, Chengquan, et al. A novel greenhouse-based system for the detection and
373 plumpness assessment of strawberry using an improved deep learning
374 technique. *Frontiers in Plant Science*. **11**,559.
375 [https://doi.org/10.3389/fpls.2020.00559\(2020\)](https://doi.org/10.3389/fpls.2020.00559(2020)).
- 376 23. Li, Yang, et al. Detection of typical obstacles in orchards based on deep
377 convolutional neural network. *Computers and Electronics in
378 Agriculture*. **181**,105932. [https://doi.org/10.1016/j.compag.2020.105932\(2021\)](https://doi.org/10.1016/j.compag.2020.105932(2021)).
- 379 24. Zhang, Kaihua, and Haikuo Shen. Solder Joint Defect Detection in the Connectors
380 Using Improved Faster-RCNN Algorithm. *Applied Sciences*. **11**,576.
381 [https://doi.org/10.3390/app11020576\(2021\)](https://doi.org/10.3390/app11020576(2021)).
- 382 25. Jiang, Du, et al. Semantic segmentation for multiscale target based on object
383 recognition using the improved Faster-RCNN model. *Future Generation Computer
384 Systems*. **123**,94-104. [https://doi.org/10.1016/j.future.2021.04.019\(2021\)](https://doi.org/10.1016/j.future.2021.04.019(2021)).
- 385 26. Kala, Jules Raymond, et al. Assets management on electrical grid using Faster-
386 RCNN. *Annals of Operations Research*. 1-14.[https://doi.org/10.1007/s10479-020-
387 03650-4\(2020\)](https://doi.org/10.1007/s10479-020-03650-4(2020)).
- 388 27. Shashidhar, R., and Sudarshan Patilkulkarni. Visual speech recognition for small
389 scale dataset using VGG16 convolution neural network. *Multimedia Tools and
390 Applications*. 1-12. [https://doi.org/10.1007/s11042-021-11119-0\(2021\)](https://doi.org/10.1007/s11042-021-11119-0(2021)).
- 391 28. Zan, Xuli, et al. Automatic Detection of Maize Tassels from UAV Images by
392 Combining Random Forest Classifier and VGG16. *Remote Sensing*. **12**,3049.
393 [https://doi.org/10.3390/rs12183049\(2020\)](https://doi.org/10.3390/rs12183049(2020)).
- 394 29. He, Kaiming, et al. Deep residual learning for image recognition. *Proceedings of*

- 395 *the IEEE conference on computer vision and pattern recognition.* (2016).
- 396 30. Zanchetta, Andrea, and Stefano Zecchetto. Wind direction retrieval from Sentinel-
397 1 SAR images using ResNet. *Remote Sensing of Environment.* **253**,112178.
398 <https://doi.org/10.1016/j.rse.2020.112178>(2021).
- 399 31. Wang, Xun, et al. Anthropometric Landmarks Extraction and Dimensions
400 Measurement Based on ResNet. *Symmetry.* **12**,1997.
401 <https://doi.org/10.3390/sym12121997>(2020).
- 402 32. Otsu, Nobuyuki. "A threshold selection method from gray-level histograms." *IEEE*
403 *transactions on systems, man, and cybernetics.***9**,62-66.(**1979**).
- 404
- 405
- 406
- 407

408 **List of figure caption:**

409

410

411 **Figure 1.** An example image of the egg mass of *spodoptera frugiperda*.

412

413 **Figure 2.** Comparison of example image before and after data augmentation: (a)

414 original image;(b) Image flipped horizontally;(c) Image flipped up and down;(d)

415 Image rotated 180 degrees counterclockwise;(e) Image rotated 90 degrees

416 counterclockwise;(f) transposed image.

417

418 **Figure 3.** An example of labeled egg mass image.

419

420 **Figure 4.** Structural diagram of Faster-RCNN.

421

422 **Figure 5.** Structure diagram of VGG16.

423

424 **Figure 6.** The algorithm structure of Region Proposal Networks.

425

426 **Figure 7.** Structure diagram of two types of residual units;(a) 2-layer residual

427 units;(b) 3-layer residual units.

428

429 **Figure 8.** An example image used to calculate the area.

430

431 **Figure 9.** Loss values of the two Faster-RCNN models;(a) training set loss;(b)

432 verification set loss.

433

434 **Figure 10.** Precision-recall curves of the two Faster-RCNN models;(a) the original

435 model;(b) the improved model.

436

437 **Figure 11.** Sample test results of the improved model.

438

439 **Figure 12.** An example of original image and image intercepted from output box;(a)

440 original image;(b) image intercepted from output box.

441

442 **Figure 13.** An example of grayscale image and binarization image of output box;(a)

443 grayscale image of output box;(b) binarization image of output box

444

445 **Figure 14.** An example of the binary image of the original image

446

Figures

Figure 1

An example image of the egg mass of *spodoptera frugiperda*.

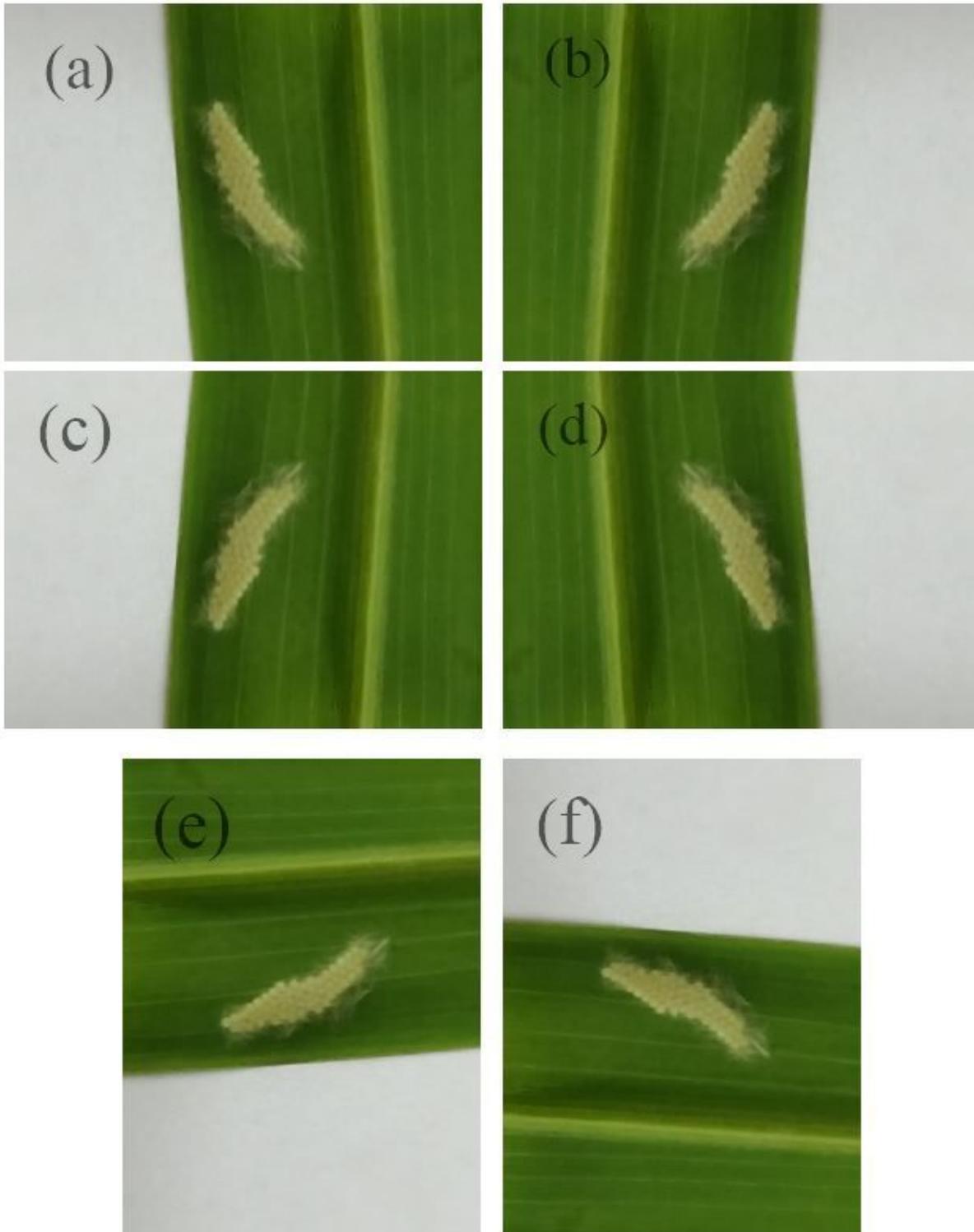


Figure 2

Comparison of example image before and after data augmentation: (a) original image;(b) Image flipped horizontally;(c) Image flipped up and down;(d) Image rotated 180 degrees counterclockwise;(e) Image rotated 90 degrees counterclockwise;(f) transposed image.

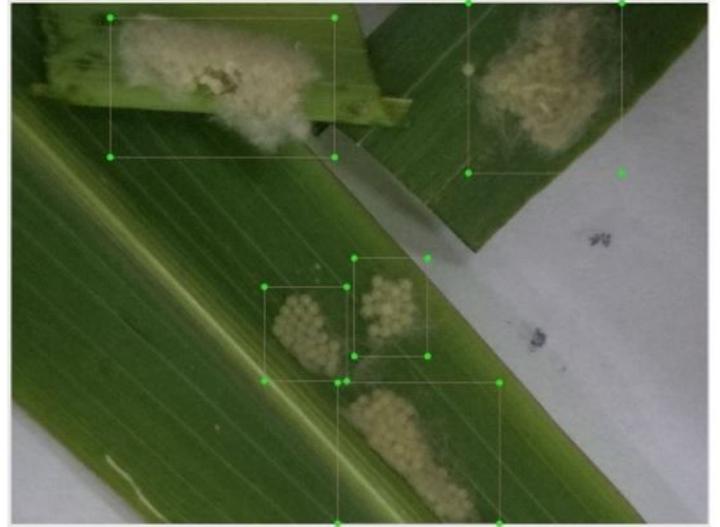


Figure 3

An example of labeled egg mass image.

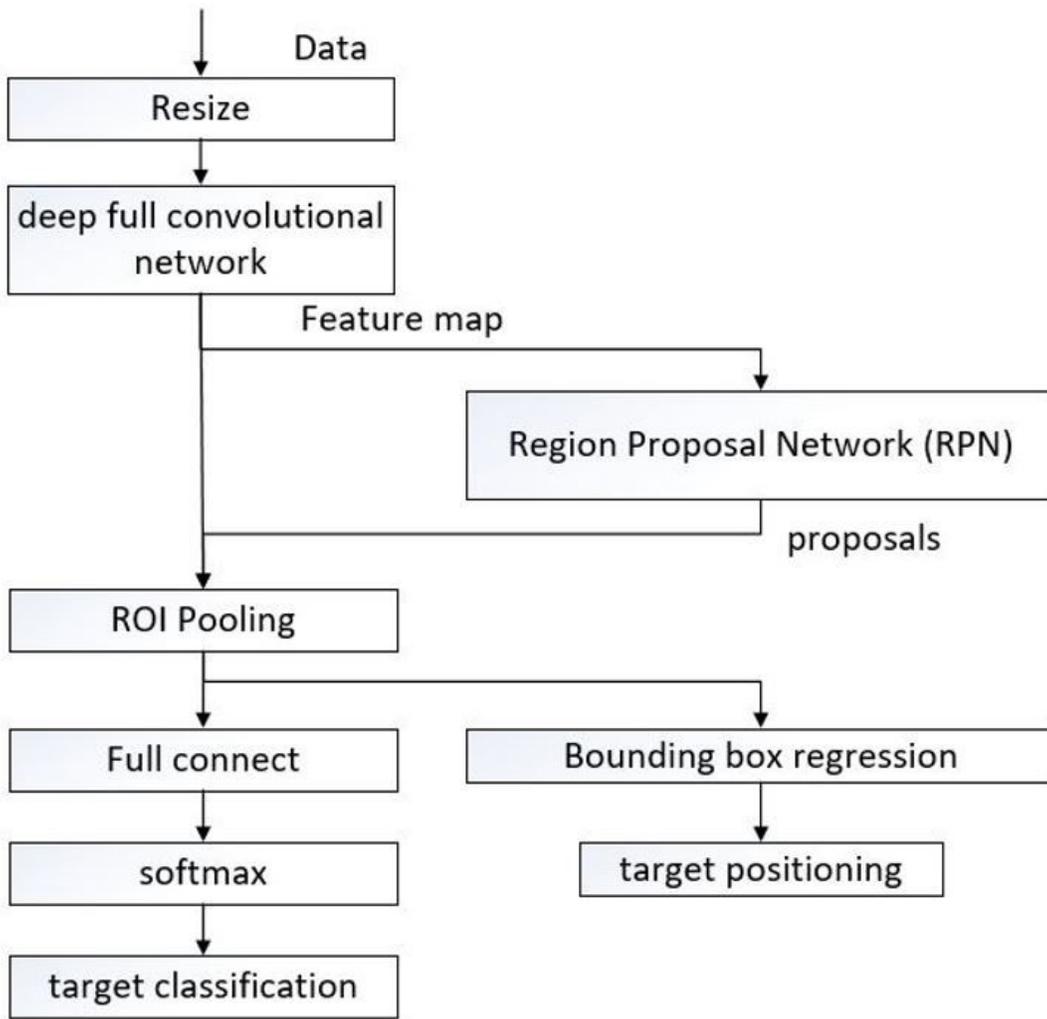


Figure 4

Structural diagram of Faster-RCNN.

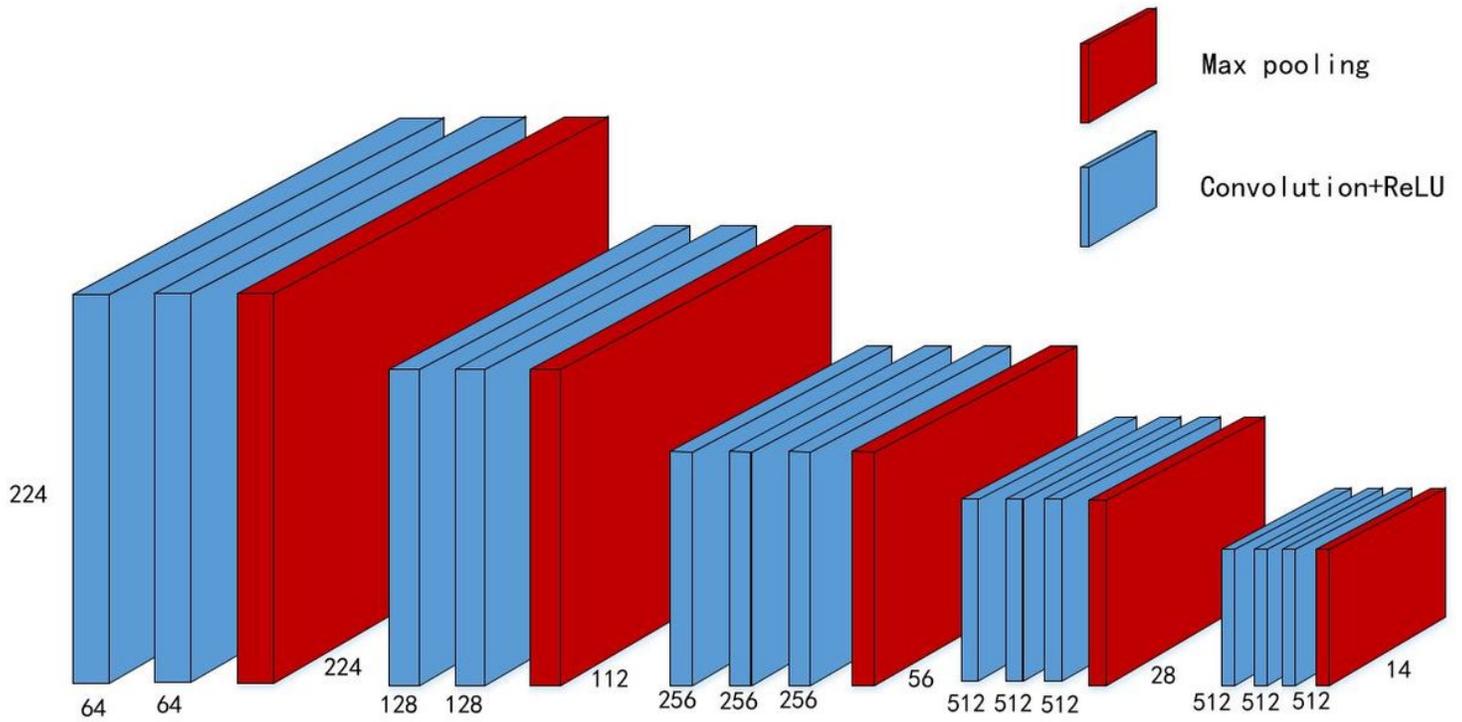


Figure 5

Structure diagram of VGG16.

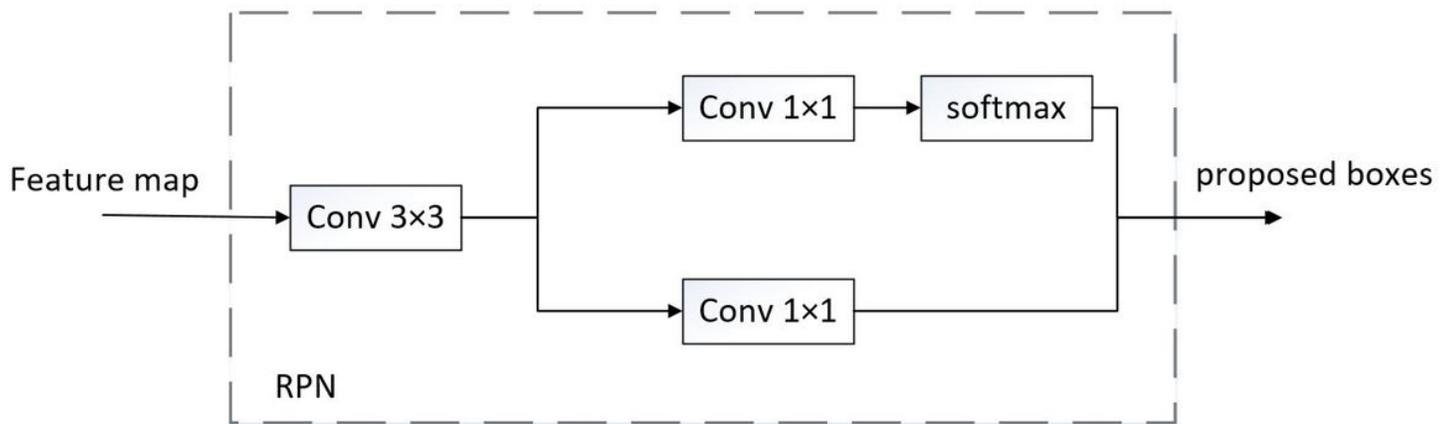


Figure 6

The algorithm structure of Region Proposal Networks.

Figure 7

Structure diagram of two types of residual units;(a) 2-layer residual units;(b) 3-layer residual units.

Figure 8

An example image used to calculate the area.

Figure 9

Loss values of the two Faster-RCNN models;(a) training set loss;(b) verification set loss.

Figure 10

Precision-recall curves of the two Faster-RCNN models;(a) the original model;(b) the improved model.

Figure 11

Sample test results of the improved model.

Figure 12

An example of original image and image intercepted from output box;(a) original image;(b) image intercepted from output box.

Figure 13

An example of grayscale image and binarization image of output box;(a) grayscale image of output box; (b) binarization image of output box

Figure 14

An example of the binary image of the original image

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1.jpg](#)