Ukrainian Journal of Physical Optics
ISSUE 2
2024

NONDESTRUCTIVE TESTING OF DUCK EGGS DURING INCUBATION USING YOLO-LITE

GUOJUN DENG [1], JIALIANG GUO [1,3] AND QINGXU LI [2]

[1] Research Department for New-Display Technologies, Jihua Laboratory, Foshan 528251, China
[2] Department of Computer Science and Technology, Anhui University of Finance and Economics, Bengbu 233030, China, e-mail: 120220059@aufe.edu.cn
[3] School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China

Received: 06.11.2023

Abstract. At present, elimination of eggs with dead embryos in China's poultry-egg incubation industry mainly relies on manual inspection and a lack of relevant automatic-detection equipment restricts seriously the development of egg-incubation industry. Accurate identification of dead eggs is a key technical task for solving this problem. In this study we design a set of image-acquisition devices for analyzing duck eggs on the stage when they are removed from the incubation tray (25 days after incubation beginning). We suggest an improved object-detection algorithm based on a YOLOv4 network and identify dead embryos with high accuracy during the egg-hatching period. According to the characteristics of collected images of the breeding-duck eggs, we remove the head subnetwork for detecting small objects in the YOLOv4 network and simplify the backbone subnetwork of this network to improve the detection efficiency. The experimental results testify that the average recognition accuracy of our YOLO-Lite network is equal to 98.33%, the recall rate amounts to 94.12% and the single-frame image-recognition time is about 15 ms. These figures are better than the corresponding parameters 96.67%, 88.89% and 32 ms, which are typical for the technique before our improvement. Therefore our results can provide a basis for the further research and development of appropriate intelligent detection equipment.

Keywords: YOLOv4 network, image recognition, deep learning, duck eggs, nondestructive testing
UDC: 535.3, 635.5, 004.9
DOI: 10.3116/16091833/Ukr.J.Phys.Opt.2024.02021

1. Introduction

During incubation process, the development of embryos in fertilized eggs can cease due to either fluctuations in temperature and humidity or bacterial infections, which can be conventionally referred to as 'dead eggs'. The occurrence of dead eggs accompanies the entire egg-incubation process. In the hatchery industry, manual removal of dead eggs from the incubation tray is traditionally performed. This process is slow and relies heavily on human expertise.

Currently, the hatchery industry is in urgent need of automated detection systems in order to replace manual removal of dead eggs [1]. When addressing this issue and detecting dead eggs, the researchers have explored various technologies such as acoustic methods [2], infrared spectroscopy [3], thermal imaging [4], electrical characteristics [5] and hyperspectral imaging [6]. These approaches face a number of challenges associated with their cost, efficiency and stability. This makes their practical application in the production setting limited, so that most of the approaches still remain in their laboratory phase.

T. Zhu et al. [7] and Ke Sun et al. [8] have suggested rapid and efficient machine-vision techniques for non-destructive inspection of fertilized eggs. J. Zhou et al. [9] have proposed new lightweight detection architecture based on a YOLOX-Tiny framework to identify the hatching

characteristics of duck eggs during 5 first days of their hatching. In a test set consisting of 326 duck-egg images, the mean average precision of the method has been equal to 99.74%. Q. Li et al. [10] have used YOLO to detect the conditions of embryos in the duck eggs incubated for the 10 days. The appropriate results demonstrate that a MobileOne-YOLO technique achieves the detection accuracy 98.10%, the recall 98.55%, the mean average precision (at the level of 50) 97.79%, the precision 98.55%, and the frame-per-second parameter amounting to 142.8 for the images of test set. The above work serves as a technical reference for the studies presented below.

Note that the works mentioned above have been mainly focused on the duck eggs hatched for only a few days or the chicken eggs. There have been no documented researches addressing the non-destructive detection of dead embryos in the duck eggs incubated for a period of 25 days, which make use of any optical technologies. Duck eggs are characterized by a high degree of surface staining and significant variations in shell thickness, which poses notable detection problems if compared to the situation occurring with chicken eggs [11]. Generally speaking, the longer the incubation period, the more difficult is detection of the conditions of a duck embryo. Therefore, searching for the fast and efficient methods for detecting dead embryos in the duck eggs incubated during the period of 25 days has a significant practical importance. A key to a successful non-destructive detection of dead embryos in the duck eggs during their incubation lies in a rational design of an image-acquisition system, which should grasp efficiently the features of dead embryos. In this study, we design an image-acquisition system for diffused-light imaging of the duck eggs obtained in a manual-hatching process which is commonly used in the practical production. This system can simultaneously capture the images carrying the information on embryo activity in an egg for three different duck eggs, which is aligned with real-world production processes.

Our analysis involves simultaneous imaging of three different eggs and solving the appropriate object-detection problem. Two approaches are typically used in the field: manual extraction of image features and automatic feature extraction based on deep learning. The former is time-consuming and labour-intensive and, moreover, it often yields suboptimal detection results. The object-detection algorithms based on deep learning are known for their high efficiency and speed. They have already found mature applications in the agricultural sector. For example, Y. Tian et al. [12] have utilized a “you only look once” (YOLO) method for detecting apples at different stages of their growth in orchards, while R. Gai et al. [13] have employed the YOLO approach for rapid detection of cherries. C. Yu et al. [14] have applied a mask region-based convolutional neural network (Mask R-CNN) for fish-morphology feature segmentation and measurements, while Z. Hao et al. [15] have used Mask R-CNN for automated detection of tree crowns and heights in young artificial forests. Moreover, Q. Liang et al. [16] have employed a single-shot multibox detector (SSD) network for the real-time detection of mango fruits on trees.

Our present work addresses the issue of removing the duck eggs with dead embryos from the incubation tray, thus providing a technical support for the development of relevant automated equipment. Therefore both the speed and the accuracy of the analysis of input images of the duck eggs are of high importance. To achieve this point, we conduct a comprehensive comparison of two of the best-performing single-step object-detection algorithms, SSD-MobileNetV3 [17] and YOLOv4 [18], taking into consideration their detection speed and accuracy. Given a superior performance of the YOLOv4 network, we choose it for detecting dead embryos in the duck eggs. Simultaneously, we plan to remove a head subnetwork responsible for detecting small targets and simplify a backbone subnetwork in the YOLOv4 network, thus making it more lightweight. The

improvement in the network performance further enhances the speed and the accuracy of dead-embryo detection, which can aid in developing the corresponding real-life equipment.

2. Materials and Methods

2.1. Collection of Activity-Information Images for Duck Eggs

A total of 606 fertilized eggs from Guoshao No. 1 Muscovy ducks originated from Shangrao, Jiangxi, were selected for our study. These eggs were disinfected by wiping the surface with alcohol and individually numbered. Subsequently, they were placed in an intelligent incubator for incubation. Every 24 hours, 10 randomly selected duck eggs were placed in a refrigerator and cold-stored for 12 hours to simulate the preparation of 'dead embryos.' After cold storage, these eggs were returned to the incubator.

An image-acquisition system was designed to mimic real-world production processes. The system consisted of a CCD camera, a light source, a dark box and a computer, as shown in Fig. 1. Images were captured at 24-hour intervals. In total, 202 images of the duck eggs were collected, with each image containing three duck eggs.

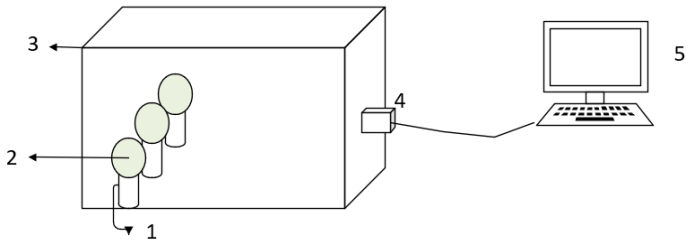
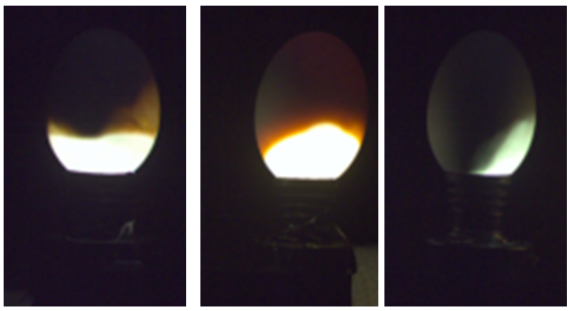


Fig. 1. Schematic diagram of our image-acquisition system: 1 – light source, 2 – duck eggs, 3 – dark box, 4 – camera and lens, and 5 – computer.

On the 26th day of incubation, the duck eggs were subjected to shell-cracking treatment in order to observe whether the duck embryos had survived. This served as an authentic basis for subsequent evaluation of models.

Fig. 2 shows the original images of duck eggs, which contain the information on embryo activity. They are taken after the eggs are removed from the incubation tray on the 25th day. These images have the resolution of 1634 pixels × 1234 pixels. Since the deceased embryos have ceased their development, such substances as egg white, yolk and blood within an egg deteriorate gradually in the high-temperature environment [19]. When these eggs are illuminated from their blunt end with a LED light (5 W cold-white), the images appear to be yellow (see Fig. 2b) or dark yellow (see Fig. 2a). This colour difference is significant in comparison with the normal embryos of which images are green (see Fig. 2c). Moreover, the difference increases with increasing time passed since the embryo death. This phenomenon enables machine-vision technologies for non-destructive detection of dead and normal duck embryos.



(a) (b) (c)

Fig. 2. Original images of duck eggs: (a) dark yellow (dead embryo), (b) yellow (dead embryo), and (c) green (normal embryo).

2.2. Dataset Preparation

Deep learning often requires a substantial amount of image data to achieve satisfactory detection results. For this purpose, a dataset of 202 images has been divided into a training set (182 images) and a test set (20 images). Data-augmentation techniques such as flipping and translation have been applied to the training set for creating an augmented dataset. This has resulted in a total of 460 images in the training set, which include 1440 duck-egg images.

After data augmentation, the images have been re-divided into the training, validation and test sets. A new set configuration includes 420 images in the training set, 40 images in the validation set and original 20 images in the test set. When employing object-detection algorithms for the duck-egg detection, manual labelling of the duck-egg images is necessary. We have employed a LabelImg annotation tool to annotate the images of the duck eggs. It marks the positions and the categories of the dead and normal duck embryos. Then the annotated data has been stored.

3. Network for Detecting Duck-Egg Activity Information

3.1. Structure of SSD-MobileNetV3 Network

An SSD network is a single-stage object-detection algorithm introduced by W. Liu et al. [20] in 2016. It has been widely applied in engineering applications with continuous improvements (see Ref. [21]). The SSD network comprises base and auxiliary subnetworks. The base network typically consists of a convolutional neural network for automatic feature extraction, while the auxiliary network is used to generate fixed-sized and scaled bounding boxes. Finally, it combines a non-maximum suppression to predict the target positions and categories. The MobileNetV3 network [22] incorporates efficiently the advantages of depth-wise separable convolutions and an inverse residual structure (IRS) from MobileNetV1 and MobileNetV2. It enhances detection speed and accuracy by introducing an 'attention mechanism', improved IRS modules and an H-Swish activation function. The MobileNetV3 network has two versions, large and small ones. MobileNetV3_large comprises 19 layers, while MobileNetV3_small consists of 15 layers. MobileNetV3_large offers slightly higher detection accuracy but slower detection speed if compared to the small version. Since we aim to detect reliably the dead and normal embryos in the duck eggs incubated for 25 days, with plans for further practical deployment, both the detection speed and the accuracy are crucial. Therefore, we choose the MobileNetV3_large network and replace the base network in the SSD for detecting and recognizing the dead and normal embryos.

In the SSD-MobileNetV3 network structure, the base subnetwork of the original SSD network is replaced with a MobileNetV3 network with no output and classification layers. Additionally, it utilizes an inverse residual structure to replace a standard convolution in the original-SSD auxiliary network, thus making the network more lightweight. This results in higher detection speed and accuracy for the SSD-MobileNetV3 network. The structure of the SSD-MobileNetV3 network is illustrated in Fig. 3. The 14th layer of the MobileNetV3 network outputs 20×20×3 anchor boxes to the detection framework, while the output of the 17th layer is passed to the first layer of the SSD auxiliary network (IRS Conv_1) which generates 10×10×6 anchor boxes. Similarly, the outputs of the second, third, fourth and fifth layers of the SSD auxiliary network (IRS Conv_2, IRS Conv_3, IRS Conv_4 and IRS Conv_5) pass to the detection framework, producing 5×5×6, 3×3×6, 2×2×6 and 1×1×6 anchor boxes, with a total of 2034 anchor boxes. By computing the Intersection over Union (IoU) between the anchor boxes and the true boxes, one removes some redundant anchor boxes to obtain prior boxes. Then, based on the location

information of the true boxes, the prior boxes are adjusted to obtain more accurate target-position information. A softmax function is used to predict the category information for each prior box. Finally, a non-maximum suppression algorithm is applied to obtain the detection results.

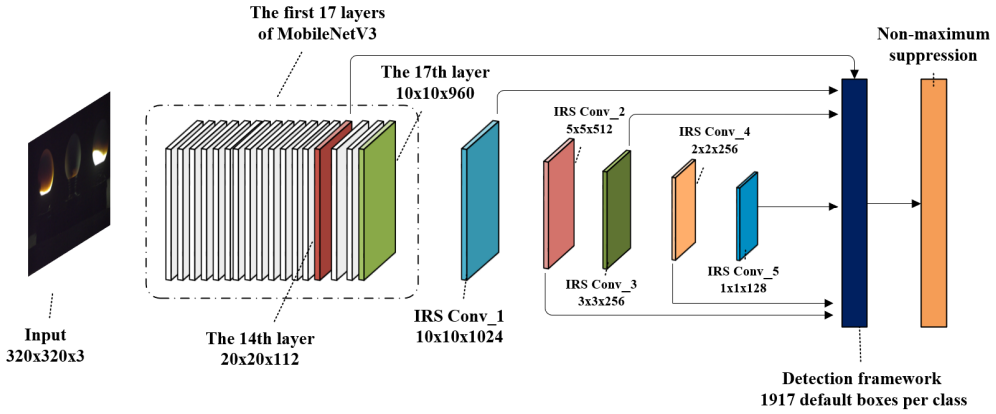


Fig. 3. SSD-MobileNet v3 architecture.

The training target-detection networks like SSD-MobileNetV3 differ from those used in image-classification tasks as they require calculating both the classification and localization losses. The goal of the network training is to reduce both the classification and localization losses to a minimal level, thus ensuring accurate prediction of the target positions and categories. The loss function for SSD-MobileNetV3 combines a weighted sum of the localization and classification losses:

$$L(x,c,l,g)=\frac{1}{N}\left[L_{conf}(x,c)+\alpha L_{loc}(x,l,g)\right]. \tag{1}$$

The following notation is used in Eq. (1):

$$L_{conf}(x,c)=-\sum_{i\in p_{os}}^{N}x_{ij}^{p}\log\hat{c}_i^{p}-\sum_{i\in n_{eg}}\log\hat{c}_i^{0}, \tag{2}$$

$$L_{loc}(x,l,g)=\sum_{i\in p_{os}}^{N}\sum_{m\in b_{ox}}x_{ij}^{p}S_{L1}\left[l_i^{t}\left(l_i^{t}-\hat{g}_i^{t}\right)\right]\quad x_{ij}^{p}\in(0,1), \tag{3}$$

$$\hat{c}_i^{p}=\frac{\exp\left(c_i^{p}\right)}{\sum_{p}\exp\left(c_i^{p}\right)}. \tag{4}$$

Here N is the number of matched default boxes, l the predicted bounding-box coordinates, g the true bounding-box coordinates, c the confidence scores for each category calculated using the Softmax function, x implies the matching flag between the true and predicted bounding boxes, S_{L1} the smooth $L1$ loss between the predicted and true positions, α the weighting coefficient, L_{loc} the position loss, L_{conf} the confidence loss, x_{ij}^{p} the matching flag for the i-th predicted box and the j-th true box of a category p, p_{os} and n_{eg} denote respectively the numbers of positives and negatives in the sample, b_{ox} is the collection of the centre coordinates and dimensions of the predicted boxes, c_i^{p} the confidence score of a category p for the i-th predicted box, $\hat{c}_i^{p}$ denotes the probability that the i-th predicted box contains a category p, $\hat{c}_i^{0}$ the probability that the i-th predicted box does not contain any objects, l_i^{t} is the i-th predicted box (with t being the centre coordinates and the dimensions of the predicted box), and $\hat{g}_i^{t}$ denotes the position information of the i-th true box.

3.2. YOLOv4 Network Architecture

The YOLO series of algorithms represents mature single-stage object-detection methods. Among them, YOLOv3 is known for its higher detection speed and accuracy. It has already found wide applications in industrial [23] and agricultural [24] domains. As seen from Fig. 4, the network structure of YOLOv3 employs the Darknet53 network [25] for feature extracting, which excludes fully connected layers. This feature extractor consists of a single convolutional layer (Conv2d) and five residual modules (Res1 to Res5). The detection process of YOLOv3 proceeds as follows. The duck-egg images are input into the feature-extraction network, with three features on different scales fed onto the feature-pyramid network. The Res5 residual module produces the features with the dimensions 13×13×1024. They are followed by five convolutional layers, resulting in 13×13×1024 features. One branch passes through two convolutional layers, leading to 13×13×21 feature maps. Through convolution and deconvolution operations, another branch concatenates with the output features of the Res4 residual module, thus producing 26×26×768 features. Following five convolutional layers, it yields 26×26×256 features. One branch undergoes two convolutional layers and results in 26×26×21 feature maps. Another branch, through convolution and deconvolution operations, concatenates with the output features of the Res3 residual module and yields in 52×52×384 features. Finally, it outputs 52×52×21 feature maps after two convolutional layers. The features on three different scales divide the original image into the grids with different sizes, e.g., a 13×13×21 feature map partitions the image into 13×13 grids. Each grid generates three anchor boxes at a fixed scale. Hence, the final output includes four

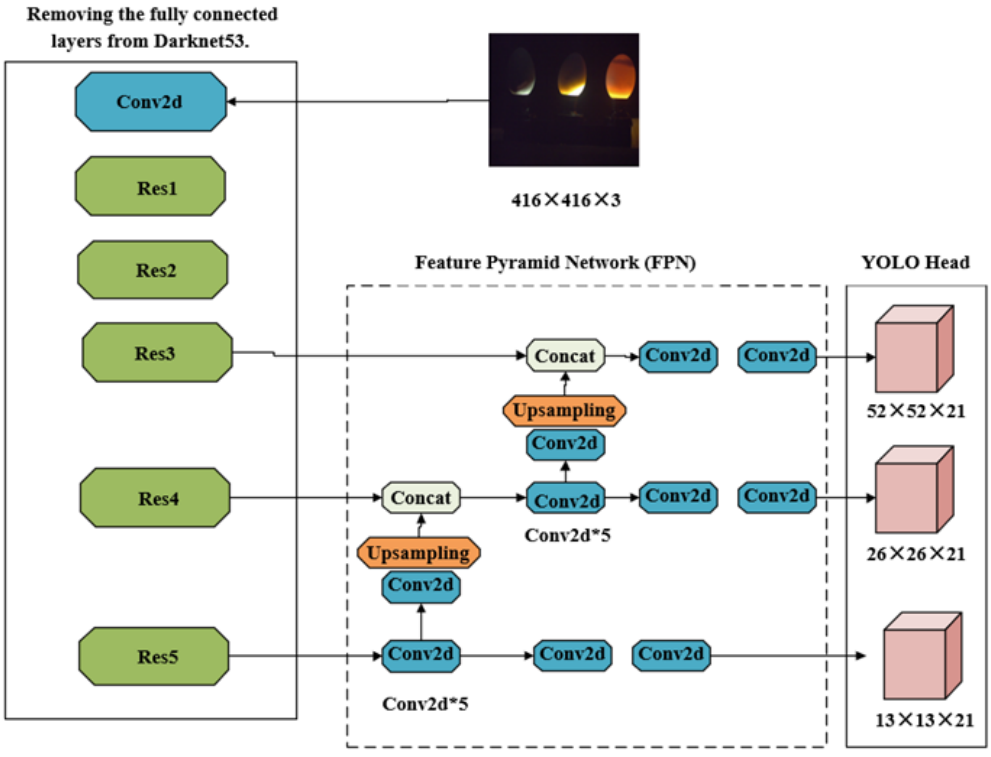


Fig. 4. YOLOv3 architecture.

coordinate values for each anchor box, one confidence score (indicating whether the anchor box contains an object) and two class scores (distinguishing between the dead and normal embryos). In total, we have 21 pieces of information. During training, the target location and the confidence information are adjusted, while a logistic function is used to predict the class information. This ultimately accomplishes the object recognition.

YOLOv4 is an improved version of YOLOv3, with significant enhancements in the input, the feature-extraction networks and the loss functions, which results in improved speed and accuracy. It employs a Mosaic algorithm for data augmentation, which selects randomly four images from the input image set and performs the operations like scaling, cropping and distortion before concatenating them. Here Fig. 5 shows four images randomly selected by YOLOv4. This enhances a generalization capability of the model. Additionally, YOLOv4 replaces the Darknet53 network used in YOLOv3 with the CSPDarknet53 network as a feature-extraction network. The CSPDarknet53 network optimizes the residual modules in the Darknet53 network, using a cross-stage partial network (CSP), and so enhances the learning capacity of the convolutional neural network while keeping it lightweight (see Fig. 6). Furthermore, YOLOv4 replaces a leaky ReLU activation function in Darknet53 with a Mish activation function:

$$\mathrm{Mish}(x) = x * \tanh\left(\ln\left(1 + e^{x}\right)\right). \tag{5}$$

Smoothness and non-monotonicity of the Mish function update efficiently the parameters of most of neurons and still preserve all of its optimization capabilities.

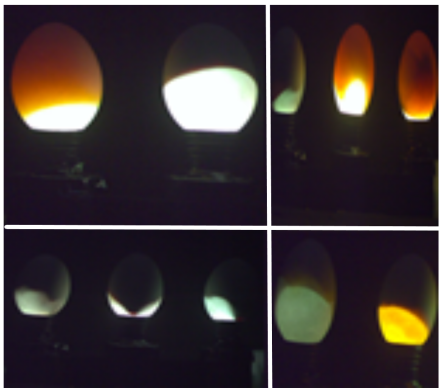


Fig. 5. Four images randomly selected by YOLOv4.

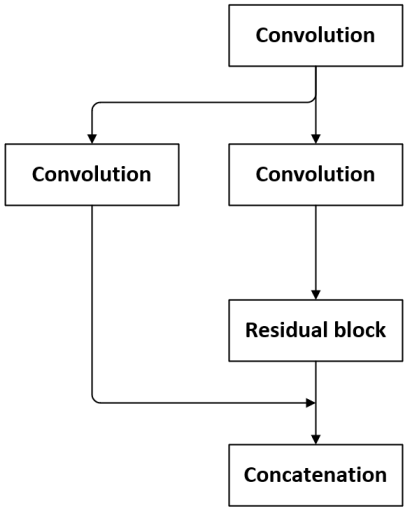


Fig. 6. CSP structure.

YOLOv4 increases significantly the receptive field of an image by introducing a spatial-pyramid pooling structure, which improves the feature-extraction capability of the network [26]. It utilizes a known path-aggregation network instead of a feature-pyramid network for feature fusion and extraction. After implementing the feature extraction from top to bottom, it proceeds with a bottom-up feature up-sampling to prevent the information loss. Moreover, YOLOv4 utilizes a complete-IoU (CIoU) procedure when computing regression localization loss. CIoU calculates the similarity between the centre distance, the overlapping area and the aspect ratio of the real box and the predicted box:

$$\mathrm{CIou} = \mathrm{IoU} - \left[\frac{\rho^2(b, b^{gt})}{d^2} + \alpha v\right], \tag{6}$$

with

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right), \tag{7}$$

$$\alpha = \frac{v}{1 - \mathrm{IoU} + v}. \tag{8}$$

Here IoU denotes the intersection-over-union between the real and predicted boxes, ρ the Euclidean distance between the centres of the predicted and real boxes, b the central point of the predicted box, b^{gt} the central point of the real box, d the Euclidean distance between the diagonal of the predicted and real boxes, v the similarity in the aspect ratio between the predicted bounding box and the ground-truth box, α the weighting coefficient, w^{gt} the width of the ground-truth box, h^{gt} the height of the ground-truth box, and w and h are respectively the width and the height of the predicted bounding box.

The YOLOv4 architecture is illustrated in Fig. 7. The duck-egg image passes through a CBM layer (convolution + batch normalization + Mish activation function) and then goes to a CSP1 layer consisting of CBM and one residual component (see Fig. 8). Subsequently, it enters a CSP2 layer comprising CBM and two residual components, two CSP8 layers comprising CBM and eight residual components, a CSP4 layer comprising CBM and four residual components, and a CBL layer (convolution + vatch normalization + leaky ReLU activation function). Afterward, the input passes through a spatial-pyramid pooling structure and a path-aggregation network, thus outputting the features on the three different scales (52×52×21, 26×26×21 and 13×13×21) to a YOLO head. Recognition and detection of the information concerned with the duck-egg condition in YOLOv4 is performed similar to YOLOv3.

Training YOLOv4 involves an iterative process of correcting the network predictions for the position, the confidence and the class information, which makes the predictions of the true values. Consequently, YOLOv4 training requires simultaneous computation of the position loss, the confidence loss and the classification loss:

$$Loss = L_{\mathrm{CIoU}} + L_{conf} + L_{class}, \tag{9}$$

with the notation

$$L_{\mathrm{CIoU}} = 1 - \mathrm{CIoU}, \tag{10}$$

$$\begin{aligned} L_{conf} = &-\sum_{i=0}^{K\times K}\sum_{j=0}^{M} I_{ij}^{obj}[\hat{C}_i \log(C_i) + (1-\hat{C}_i)\log(1-C_i)] \\ &-\sum_{i=0}^{K\times K}\sum_{j=0}^{M} I_{ij}^{noobj}[\hat{C}_i \log(C_i) + (1-\hat{C}_i)\log(1-\hat{C}_i)] \end{aligned}, \tag{11}$$

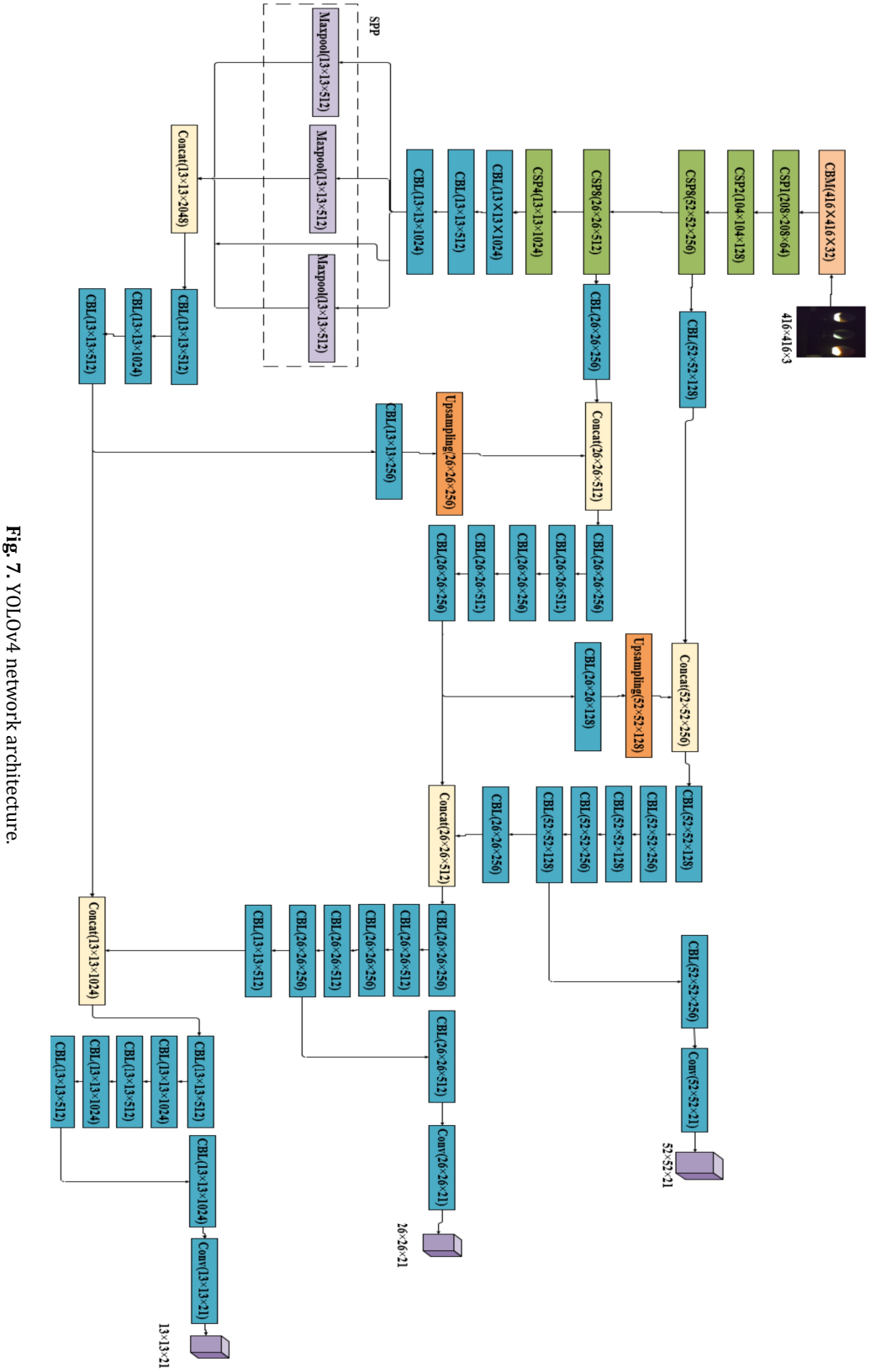


Fig. 7. YOLOv4 network architecture.

$$L_{class} = -\sum_{i=0}^{K\times K}\sum_{j=0}^{M} I_{ij}^{obj} \sum_{c\in classes} \left\{ p_i(c)\log\left[\hat{p}_i(c)\right] + \left[1-p_i(c)\right]\log\left[\hat{p}_i(c)\right] \right\}. \quad (12)$$

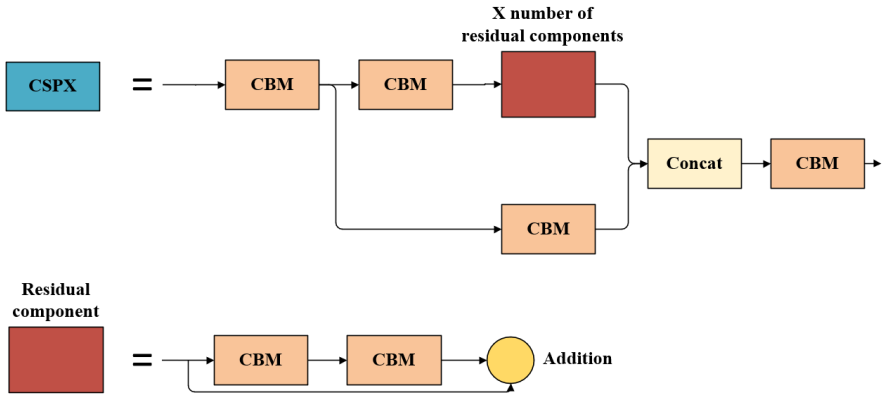


Fig. 8. CSP1 layer architecture.

In Eqs. (9)–(12), L_{CIoU} means the regression loss, L_{class} the classification loss, the CIoU parameter is given by Eq. (6), K is the number of grid divisions, M the number of prior boxes generated per grid, I_{ij}^{obj} signifies whether the j-th prior box in the i-th grid is responsible for predicting the target (with 1 if it is responsible and 0 otherwise), I_{ij}^{noobj} implies that the j-th prior box in the i-th grid is not responsible for predicting the target, $\hat{C}_i$ is the confidence score of the predicted bounding box, C_i the confidence score of the ground-truth bounding box, $\mathtt{classes}$ denotes the set of categories for classification, $p_i(c)$ the ground-truth category of the target, and $\hat{p}_i(c)$ the predicted category of the target-bounding box.

3.3. Improved YOLOv4 Network Architecture

Although YOLOv4 demonstrates significant advantages in detecting small targets, we remind that the primary objective of our study is to achieve faster and more accurate detection of dead and normal embryos in the duck eggs during their incubation. It is noteworthy that, in their original images, the duck eggs can hardly be categorized as small targets. Since the 52×52×21 YOLO head in the YOLOv4 network is designed primarily for detecting small-scale targets, it has been removed in this study. Although the spatial-pyramid pooling structure enhances the feature-extraction capabilities of the network, it also increases the number of the parameters and so affects the network speed to some extent. In order to enhance the network speed in distinguishing dead and normal embryos, we eliminate the spatial-pyramid pooling structure.

The feature-extraction network in YOLOv4 is based on CSPDarknet53, which has deep-network architecture and results in a slower detection speed. To address this issue, we have conducted a number of repeated experiments and their validation. As a consequence, we have constructed a convolutional neural network optimized due to the CSP structure and the Mish activation function in order to extract automatically the vital information features from the duck eggs.

As illustrated in Fig. 9, the improved YOLOv4 network structure employs a mosaic algorithm to implement the data augmentation for the duck-egg images. These images are then fed into the backbone network, which comprises a CBM layer, a maxpool layer and CSPM (i.e., a convolutional multi-layer neural network formed by combining a CSP structure with CBM). The

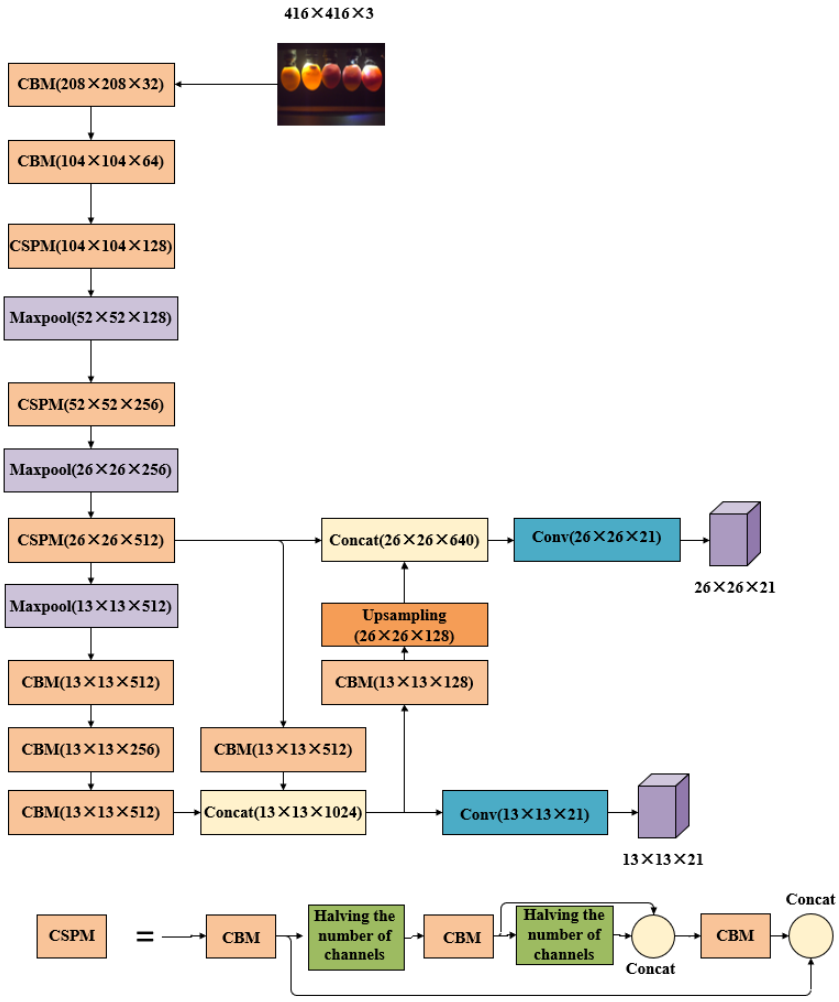


Fig. 9. Architecture of our improved YOLOv4 model.

backbone network generates the features on the two scales, 13×13 and 26×26, for detecting the medium and large-sized targets. The path-aggregation network is used to extract and fuse the features from both the lower and upper layers. This results in 13×13×21 and 26×26×21 feature maps which are fed into the YOLO head network for predicting the position and the category information for the dead and normal embryos.

4. Model Training and Analysis of Results

4.1. Experimental Platform

A hardware platform for the network training is as follows: AMD Ryzen Threadripper 2920X CPU, NVIDIA GeForce RTX 2080Ti GPU, 128 GB of RAM. Our software platform is given by Windows 10 operating system, CUDA 9.2 parallel computing framework + CUDNN v7.6 deep neural network acceleration library, open-source deep learning framework darknet based on C language and CUDA + VS2017 + Cmake 3.18, and Tensorflow-GPU 1.13.1 + Python 3.6.

4.2. Model Training

A transfer learning allows the model to converge quickly on a small dataset and so it addresses, to some extent, the issue of a limited number of samples. In this study, we have

trained the classification task for the dead and normal embryos, using pre-trained weights from a Microsoft Common Objects in Context (COCO) dataset. The initial learning rate has been set to 0.001, the batch size for the training images to 8, and the maximum number of training iterations to 8000. Fig. 10 illustrates the changes occurring with the loss function during training with the improved YOLOv4 version. It is evident that the average loss decreases rapidly well before 1000 training epochs. After 1000 epochs, it becomes stabilized at a low enough level, thus indicating that the model has already converged.

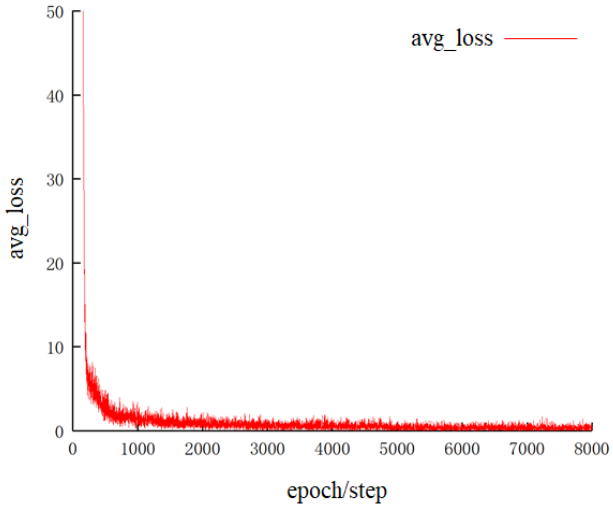


Fig. 10. Loss curve obtained for our improved YOLOv4 model used in the experimental studies during training.

4.3. Analysis of Results

4.3.1. Evaluation Metrics

To assess the performance of our object-detection algorithms, first we have evaluated the mean average precision of the model, the detection speed and the miss rate on the test dataset. A higher mean average precision indicates better predictive performance of the model. The detection speed measures the number of images the model can process per unit time, while the miss rate is a crucial metric to assess the reliability of a model, mainly reflecting the number of undetected objects. The precision and the recall can be calculated as follows:

$$\begin{cases} R = \dfrac{TP}{TP + FN} \\ P = \dfrac{TP}{TP + FP} \end{cases}. \tag{13}$$

In Eqs. (13), R denotes the recall, P the precision, TP is the number of correctly classified normal embryos as positive samples, FN the number of normal embryos misclassified as negative samples, and FP represents the number of dead embryos misclassified as positive samples.

4.3.2. Data Analysis

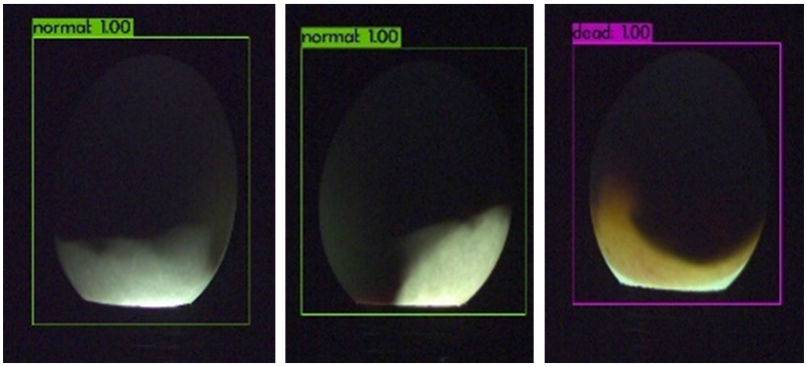
We have analyzed statistically the test results for 20 images of duck eggs taken after 25 days of their incubation, which includes a total of 60 duck eggs. The evaluation criteria consist of the accuracy, the recall and the detection speed. The main results are shown in Table 1. Both SSD-MobileNetV3 and YOLOv3 misclassify 3 duck eggs, YOLOv4 misclassify 2 eggs, and the improved YOLOv4 misclassify only 1 egg. When compared to YOLOv3 and SSD-MobileNetV3, YOLOv4 demonstrates a superior performance in terms of both the detection speed and the

accuracy. Therefore, one can choose YOLOv4 as a basis for further detection improvements. Note that the detection speed is the number of images which the model can detect per second when run on a CPU.

Table 1. Results of performance tests for detection based on different models.

Model	Accuracy, %	Recall rate, %	Detection speed, f/s
SSD-MobileNetV3	95	84.21	20
YOLOv3	95	84.21	28.57
YOLOv4	96.67	88	31.25
Improved YOLOv4	98.33	94.12	66.7

As seen from Table 1, the improved YOLOv4 network achieves very high detection accuracy and recall rates (98.33% and 94.12%, respectively) in the detection of vitality information for the duck embryos. This represents a significant improvement when compared with SSD-MobileNetV3, YOLOv3 and YOLOv4. Furthermore, our improved model achieves the detection speed 66.7 f/s, which is significantly higher than the next-to-highest result, 31.25 f/s. Hence, our approach is suitable for the real-time detection. All of the four models under test do not exhibit any instances of missed detection, which indicates their high reliability. Finally, Fig. 11 illustrates the recognition results of the improved YOLOv4 network obtained for the dead and normal duck embryos in incubated duck eggs.



(a) (b) (c)

Fig. 11. Test results for the duck eggs derived from our improved YOLOv4 model: panels (a) and (b) correspond to normal embryos, and panel (c) to dead embryo.

5. Conclusions

Let us summarize the main results obtained in the present study.

We have simplified the YOLOv4 backbone network, used the CSP structure and the Mish activation function, and have removed the YOLO head designed for detecting small objects. In this manner we have reached the recognition accuracy 98.33% and the recall rate 94.12% for distinguishing between the dead and normal duck embryos in the incubated duck eggs. When the model is run on a CPU, the detection speed is equal to 66.7 f/s. This demonstrates high robustness of our model and its suitability for practical production requirements.

By comparing the three object-detection algorithms, SSD-MobileNetV3, YOLOv3 and YOLOv4, we have found that YOLOv4 exhibits better performance in classifying the dead and normal duck embryos in the incubated duck eggs. This has led us to taking YOLOv4 as a basic model for further improvement.

We have designed an image-capture system suitable for detecting the information on duck-embryo vitality, which is suitable for the real production process. By leveraging end-to-end convolutional neural networks, we are able to distinguish between the dead and normal duck embryos, using a single duck-egg image. This provides a technical reference for using machines instead of manual labour for detecting the condition of duck embryos and offers a technical support for the development of appropriate automation equipment in the future.

Funding. This study has been supported by the Natural Science Foundation of Anhui Provincial Department of Education, China (No. 2022AH050604, KJ2021A0477) and the Key Research and Development Program Projects of Anhui Province, China (No. 202104a06020014).

Disclosures. The authors declare no conflict of interest.

References

1. Liu, Y., Xiao, D., Zhou, J., & Zhao, S. (2023). AFF-YOLOX: An improved lightweight YOLOX network to detect early hatching information of duck eggs. *Computers and Electronics in Agriculture*, *210*, 107893.
2. Lai C, Li C, Huang K and Cheng C, 2021. Duck eggshell crack detection by nondestructive sonic measurement and analysis. Sensors. 21(21): 7299.
3. Dong, J., Dong, X., Li, Y., Zhang, B., Zhao, L., Chao, K., & Tang, X. (2020). Prediction of infertile chicken eggs before hatching by the Naïve-Bayes method combined with visible near infrared transmission spectroscopy. *Spectroscopy Letters*, *53*(5), 327-336.
4. Zainuddin, Z., & Achmad, A. (2023, July). Classification of Fertile And Infertile Eggs Using Thermal Camera Image And Histogram Analysis: Technology Application In Poultry Farming Industry. In *2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)* (pp. 130-135). IEEE.
5. Shi, C., Wang, Y., Zhang, C., Yuan, J., Cheng, Y., Jia, B., & Zhu, C. (2022). Nondestructive Detection of Microcracks in Poultry Eggs Based on the Electrical Characteristics Model. *Agriculture*, *12*(8), 1137.
6. Lawrence, K. C., Smith, D. P., Windham, W. R., Heitschmidt, G. W., & Park, B. (2006, October). Egg embryo development detection with hyperspectral imaging. In *Optics for Natural Resources, Agriculture, and Foods* (Vol. 6381, pp. 234-241). SPIE.
7. Zhu, T., & Wang, Q. (2011). Non-destructive detection of Sudan dye duck eggs based on computer vision and fuzzy cluster analysis. *Afr. J. Agric. Res.*, *6*, 1177-1181.
8. Sun, K., Ma, L., Pan, L., & Tu, K. (2017). Sequenced wave signal extraction and classification algorithm for duck egg crack on-line detection. *Computers and Electronics in Agriculture*, *142*, 429-439.
9. Zhou, J., Liu, Y., Zhou, S., Chen, M., & Xiao, D. (2023). Evaluation of Duck Egg Hatching Characteristics with a Lightweight Multi-Target Detection Method. *Animals*, *13*(7), 1204.
10. Li, Q., Shao, Z., Zhou, W., Su, Q., & Wang, Q. (2023). MobileOne-YOLO: Improving the YOLOv7 network for the detection of unfertilized duck eggs and early duck embryo development-a novel approach. *Computers and Electronics in Agriculture*, *214*, 108316.
11. Dong, J., Dong, X., Li, Y., Peng, Y., Chao, K., Gao, C., & Tang, X. (2019). Identification of unfertilized duck eggs before hatching using visible/near infrared transmittance spectroscopy. *Computers and Electronics in Agriculture*, *157*, 471-478.
12. Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., & Liang, Z. (2019). Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Computers and Electronics in Agriculture*, *157*, 417-426.
13. Gai, R., Chen, N., & Yuan, H. (2023). A detection algorithm for cherry fruits based on the improved YOLO-v4 model. *Neural Computing and Applications*, *35*(19), 13895-13906.
14. Yu, C., Fan, X., Hu, Z., Xia, X., Zhao, Y., Li, R., & Bai, Y. (2020). Segmentation and measurement scheme for fish morphological features based on Mask R-CNN. *Information Processing in Agriculture*, *7*(4), 523-534.
15. Hao, Z., Lin, L., Post, C. J., Mikhailova, E. A., Li, M., Chen, Y., Yu, K., & Liu, J. (2021). Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN). *ISPRS Journal of Photogrammetry and Remote Sensing*, *178*, 112-123.
16. Liang, Q., Zhu, W., Long, J., Wang, Y., Sun, W., & Wu, W. (2018). A real-time detection framework for on-tree mango based on SSD network. In *Intelligent Robotics and Applications: 11th International Conference, ICIRA 2018, Newcastle, NSW, Australia, August 9–11, 2018, Proceedings, Part II 11* (pp. 423-436). Springer International Publishing.
17. Lu, H., Li, C., Chen, W., & Jiang, Z. (2020). A single shot multibox detector based on welding operation method for biometrics recognition in smart cities. *Pattern Recognition Letters*, *140*, 295-302.
18. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
19. Geng, L., Yan, T., Xiao, Z., Xi, J., & Li, Y. (2018). Hatching eggs classification based on deep learning. *Multimedia Tools and Applications*, *77*, 22071-22082.

20. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing.
21. Lu, X., Ji, J., Xing, Z., & Miao, Q. (2021). Attention and feature fusion SSD for remote sensing object detection. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-9.
22. Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. (2019). Searching for mobilenetv3. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 1314-1324).
23. Li, Y., Lu, Y., & Chen, J. (2021). A deep learning approach for real-time rebar counting on the construction site based on YOLOv3 detector. *Automation in Construction*, *124*, 103602.
24. Lawal, M. O. (2021). Tomato detection based on modified YOLOv3 framework. *Scientific Reports*, *11*(1), 1447.
25. Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
26. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, *37*(9), 1904-1916.

Guojun Deng, Jialiang Guo and Qingxu Li. (2024). Nondestructive Testing of Duck Eggs During Incubation Using YOLO-LITE**.** *Ukrainian Journal of Physical Optics*, *25*(2)**,** 02021 – 02035.
DOI: 10.3116/16091833/Ukr.J.Phys.Opt.2024.02021

Анотація*. Наразі усунення яєць із мертвими зародками в інкубаційній промисловості птахівництва в Китаї в основному покладається на ручну перевірку, а відсутність відповідного обладнання автоматичного виявлення серйозно обмежує розвиток цієї галузі. Точна ідентифікація мертвих яєць є ключовим технічним завданням для вирішення проблеми. У цьому дослідженні розроблено установку для одержання зображень для аналізу качиних яєць на стадії, коли їх виймають з інкубаційного лотка (25 днів після початку інкубації). Ми пропонуємо вдосконалений алгоритм виявлення об'єктів на основі мережі YOLOv4 та ідентифікуємо мертві ембріони з високою точністю під час періоду висиджування яєць. Відповідно до характеристик зображень яєць племінних качок, ми видалили головну підмережу для виявлення малих об'єктів у мережі YOLOv4 і спростили магістральну підмережу цієї мережі, щоб підвищити ефективність виявлення. Експериментальні результати засвідчують, що середня точність розпізнавання нашої мережі YOLO-Lite дорівнює 98,33%, показник повноти складає 94,12%, а час розпізнавання одного кадру зображення – близько 15 мс. Ці показники кращі за відповідні параметри 96,67%, 88,89% і 32 мс, притаманні аналогічній методиці без наших удосконалень. Тому наші результати можуть стати основою для подальших досліджень і розробки відповідного інтелектуального обладнання.*

Ключові слова*: мережа YOLOv4, розпізнавання зображень, глибоке навчання, качині яйця, безконтактний контроль*