

## Research Article

# Recognition and Classification of Broiler Droppings Based on Deep Convolutional Neural Network

Jintao Wang,<sup>1</sup> Mingxia Shen ,<sup>1</sup> Longshen Liu ,<sup>1</sup> Yi Xu,<sup>2</sup> and Cedric Okinda <sup>1</sup>

<sup>1</sup>College of Engineering, Laboratory of Modern Facility Agriculture Technology and Equipment Engineering of Jiangsu Province, Nanjing Agricultural University, Jiangsu 210031, China

<sup>2</sup>New Hope Liuhe Co., Ltd., Shandong 266100, China

Correspondence should be addressed to Mingxia Shen; [mingxia@njau.edu.cn](mailto:mingxia@njau.edu.cn)

Received 19 August 2019; Accepted 30 October 2019; Published 27 November 2019

Academic Editor: Giovanni Diraco

Copyright © 2019 Jintao Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Digestive diseases are one of the common broiler diseases that significantly affect production and animal welfare in broiler breeding. Droppings examination and observation are the most precise techniques to detect the occurrence of digestive disease infections in birds. This study proposes an automated broiler digestive disease detector based on a deep Convolutional Neural Network model to classify fine-grained abnormal broiler droppings images as normal and abnormal (shape, color, water content, and shape&water). Droppings images were collected from 10,000 25-35-day-old Ross broiler birds reared in multilayer cages with automatic droppings conveyor belts. For comparative purposes, Faster R-CNN and YOLO-V3 deep Convolutional Neural Networks were developed. The performance of YOLO-V3 was improved by optimizing the anchor box. Faster R-CNN achieved 99.1% recall and 93.3% mean average precision, while YOLO-V3 achieved 88.7% recall and 84.3% mean average precision on the testing data set. The proposed detector can provide technical support for the detection of digestive diseases in broiler production by automatically and nonintrusively recognizing and classifying chicken droppings.

## 1. Introduction

Livestock production accounts for about 50% of the global agricultural gross domestic product. Additionally, it is the source of livelihood for more than 1.3 billion people in developing countries [1]. Due to the ever-growing global population, the global consumption of animal products is projected to increase by 70% by the year 2050 [2]. To sustain the growing demand for animal-based proteins, industrial large-scale livestock keeping was introduced. Currently, over 3535 million animals are reared under extensive and intensive industrial-scale production systems, with a total annual production of 3029 million tons and 798 million tons of meat and milk, respectively [3]. The main limiting factor in large-scale livestock production is the occurrence of diseases, which has a disastrous effect on production economy, food safety, and biosecurity [4]. This study focuses on broilers, which are the main contributor to global white meat production [5].

One of the major diseases in poultry production is digestive diseases such as coccidiosis, infectious bursal disease,

paratyphoid, worms, and ulcerative enteritis, to mention a few [6]. The most common technique used to detect the occurrences of these poultry diseases is visual observation and sound distinction [7]. However, this technique is time-consuming, subjective, labor-intensive, and often fails to provide early detection [8]. With the development of computer vision systems, computerized disease diagnosis and detection of sick birds have been reported in several studies. Zhuang et al. [7] performed a skeleton analysis for early detection of sick broilers by image processing. Additionally, Zhuang and Zhang [8] reported on a sick broiler detector based on deep learning techniques. However, the analysis of chicken droppings by image processing and deep learning for sick bird detection has not been reported in any literature. Different poultry diseases have almost identical symptoms. Hence, it is difficult to diagnose a specific disease by chicken posture and sound observation. Moreover, the sound discrimination technique is highly infeasible in a group setting.

The droppings of a chicken vary in color, shape, and texture depending on the diet, time of the year, chicken species,

and health. Birds defecate more than 12 times a day [6]. Therefore, any variations can be detected in real time. Hence, droppings examination is the most efficient technique used to detect any digestive tract abnormalities [6]. Normal chicken droppings are usually in shades of brown and pretty solid with a kind of small covered white on top. However, it is also typical that droppings can take different textures and colors depending on their diet. The chicken health handbook by Damerow [6] presented the effects of common digestive diseases and feed on chicken droppings. Greenish droppings can result from intestinal worms, Marek's disease, or avian flu but can also result from diet rich in vegetables, herbs, grass, weeds, and plants of all kinds. Yellowish hue droppings can be a result of coccidiosis, intestinal worms, or a disease of the kidneys but can also result from intakes of some foods such as forsythia flowers, strawberries or tomatoes, and corn. Droppings of black color can be as a result of internal bleeding or ingestion of wood ash, charcoal, and dark berries. Liquid brown droppings can be due to infection from *E. coli* or respiratory system infections such as bronchitis. However, it can also occur due to ingestion of food with high liquid content. White or highly liquid droppings can be caused by infection of the cloaca, wind gleet, damage to the kidneys, stress, and internal disease. However, can also be due to ingestion of a larger quantity than the normal amount of water. Finally, orange or red droppings can be due to coccidiosis, lead poisoning, and swelling or inflammation of the intestinal wall [6]. In intensive production where birds feed on a standard diet, a variation in their droppings can be easily detected.

Due to the superior performance of Convolutional Neural Networks (CNN), it has become a popular technique in object recognition [9]. Several CNN network structures have already been applied in object identification and classification, such as AlexNet [10], VGG, and ResNet [11]. Convolutional Neural Networks (CNN) has been applied to solve the problem of visual object recognition in an image or video because of its superior performance. Faster R-CNN and YOLO-V3 are two kinds of CNN networks with superior performance, which have the advantages of high detection accuracy and fast detection speed [12, 13]. Faster R-CNN is an improvement of Fast R-CNN. It performs a four-step target detection, i.e., region proposal, feature extraction, classification, and regression [14]. It applies the Region Proposal Network (RPN), which improves the accuracy and speed at the same time. YOLO-V3 combines classification and location into a task recognition target with the whole picture as the region of interest. It combines the feature interaction layer of three scale prediction and uses FPN algorithm to merge the features of different scales to enhance the detection effect of small target objects. Currently, CNN has also been applied in several areas in smart agriculture. Zhuang and Zhang [8] detected the sick broiler by CNN based on the Improved Feature Fusion Single Shot Multibox Detector (IFFSSD), and Zheng et al. [15] applied SBDA-DL for the real-time detection of sow behaviors. Tian et al. [16] used YOLO-V3 to detect apples in orchards during different growth cycles.

This study introduces a normal and abnormal classification technique of broiler droppings based on image processing and deep learning for digestive disease detection. The

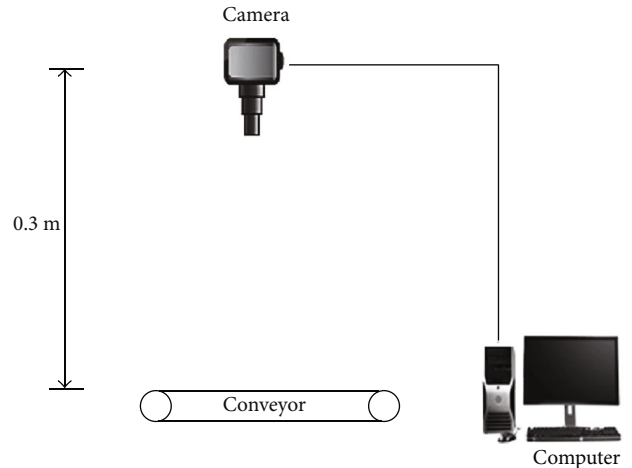


FIGURE 1: Experimental setup and image acquisition system.

main objective of this study was to develop an efficient and effective autoclassifier of healthy and infected broilers based on droppings images. The specific objectives included developing a comprehensive data set for the model training and efficiently training and developing robust deep learning models for disease detection. This proposed system would serve as a technical support tool for early warning of occurrences of digestive diseases in broilers.

## 2. Materials and Methods

### 2.1. Experimental Setup

**2.1.1. Animal Housing and Data Collection.** Experiments were conducted between March and April 2019 at Wangguancun, Jianan District, Xuchang City, Henan Province, China. A total of 10,000 25-35-day-old Ross broilers were used in the experiment. The birds were kept in a multilayer cage mode with automatic droppings conveyor belts. The broilers were fed a pellet diet of 19% protein and 6% oil and 3300 kcal/kg. Drinking water was *ad libitum* during the experiment period. The temperature and humidity were autoregulated between 22 and 24°C and 60 and 70%, respectively. The lighting regime consisted of the LED lights fixed on the ceiling and the wall period between 5:00 and 23:00 with approximately 40 lx during the light period.

A Canon 5D SLR camera (5760 by 3240 pixels) was mounted 0.3 m directly above the conveyor belt as shown in Figure 1. An additional illumination was provided during image acquisition by an LED light at 101 lx. The image acquisition time was 0900 hr to 1600 hr, depending on the established broiler's excretion habits.

**2.2. Data Labeling.** A professional veterinarian labeled the acquired images according to the description in Table 1. Figure 2 presents a description of the differences between the droppings labels.

**2.3. Image Data Augmentation.** Because of the complex structure of the deep Convolutional Neural Network composed, a large number of parameters, including weights and deviations, need to be evaluated in the training phase [9].

TABLE 1: Classification of droppings based on observation.

Label	Description
Normal	Brown, blue-gray, and white solid at the top, with a continuous regular shape
Abnormal shape	Dispersed, discontinuous, small granular, and a massive mass of droppings, for example, feed-like droppings
Abnormal color	Unusual colors, such as black, red, and green
Abnormal water content	The water content of the droppings is too high, such as liquid droppings
Abnormal shape&water	Droppings do not have the normal shape and water content and the boundary is discontinuous, usually manifested as a large piece of watery droppings. Compared with droppings with abnormal water content, it is more diluted and larger

Comparing with other simple structures, deep Convolutional Neural Networks need more training data to get the appropriate convergence. In order to achieve better recognition accuracy and to prevent data overfitting, it was necessary to preprocess the images with the ten augmentation methods given in Table 2. The image augmentation expanded the images to 10,000 images.

**2.4. Data Set Production.** The resultant images were then filtered to 7637 high-quality images based on the noise and rotation level (high noise and low rotation images were eliminated). The selected images were then compressed to 1800 by 1012 pixels to reduce computation complexity during training. The open-source software LableImage was used to perform image processing. When labeling, the continuous main body of droppings was labeled, and the fecal debris produced by spattering was ignored. Pascal VOC 2007 format was applied in this study. The summary of the data set is given in Table 3.

**2.5. The Architecture of Droppings Detection Systems.** In this study, two architectures of Deep Convolutional Neural Networks were explored for comparative analysis: firstly, the Faster R-CNN architecture and, secondly, YOLO-V3. The main difference between Faster R-CNN and YOLO-V3 is that Faster R-CNN firstly extracts the regions of interest (ROI) and then performs recognition of the image region into objects or background and refining the boundaries of these regions [12, 17]. YOLO-V3, on the other hand, used the whole image as the ROI then performs recognition. The parameters of a model greatly influence the performance of the entire network; in a deep learning network, the number of parameters depends on the network scale [17]. Based on the study by Yue et al. [18], increasing the number of parameters in higher layers boosts the performance of the deep neural models, while increasing the number of parameters in lower layers could be harmful to the performance of the models.

**2.6. Faster R-CNN Architecture.** In this research, Faster R-CNN was applied for broiler droppings detection and classification by combining the Region Proposal Network

(RPN) and Fast R-CNN [12]. In comparison to other target detection methods such as R-CNN and Fast R-CNN, Faster R-CNN improves the detection accuracy and speed, realizes end-to-end target detection, and can quickly generate candidate boxes.

This study applied RPN+ResNet models. ResNet was introduced in 2015 and won the competition of the ImageNet classification task because of its “simplicity and practicality” coexistence. ResNet is a deep residual network, which can solve the problem of decreasing accuracy in the training set with the deepening of the network [19]. It should be noted that as the number of convolution layers of the CNN network used in Faster R-CNN increases (ZF, VGG, GoogLeNet, and ResNet), the lower the detection speed. However, deeper networks may improve network accuracy [11]. In this study, the number of conv layers were 13, relu layers were 13, and pooling layers were 4.

The network structure of Faster R-CNN is shown in Figure 3. Faster R-CNN supports any size of image input, but before convolution operation, the input images are standardized or resized to the standard size [12]. The conversion of input and output sizes is given as

$$\text{output}_{\text{size}} = \frac{(\text{input}_{\text{size}}) - \text{kernel\_size} + 2 * \text{pad}}{\text{stride}} + 1 \quad (1)$$

The RPN was used to extract the ROI. The working process of RPN is a two-step operation, i.e., RPN classification and RPN bounding box regression. The process of RPN classification is a two-classification process. Firstly,  $K \times H \times W$  regions (called anchor,  $K = 9$ ,  $H$  is the height of the feature map, and  $W$  is the width) should be evenly divided on the feature map [12]. By comparing the overlap between these anchors and ground truth, we can decide which anchors are prospects and which are backgrounds, that is, label each anchor with prospects or backgrounds. With labels, RPN can be trained to recognize the foreground and background for any input [12].

A feature map had  $9 \times H \times W$  anchors; i.e., 9 anchors correspond to each point. The nine anchors have 1:1, 1:2, and 2:1 aspect ratios. Each aspect ratio has three dimensions [12]. The RPN bounding box regression is used to determine the approximate position of the foreground. After getting the approximate position of the proposal, the next step is to do the regression of the exact position. After the training convergence of RPN, we can get the offset of the anchor relative to the proposal. With the offset, the approximate position of the proposal can be calculated. Accurate positioning can be achieved by nonmaximum suppression (NMS) and other operations. In the training process of RPN, anchors with the largest intersection over union (IOU) with the ground truth box (GT box) are defined as positive samples, while negative samples are defined as IOU with a GT box less than 0.3 samples [12]. The loss function for RPN during the training is defined as

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum iL_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum i p_i^* L_{\text{reg}}(t_i, t_i^*), \quad (2)$$

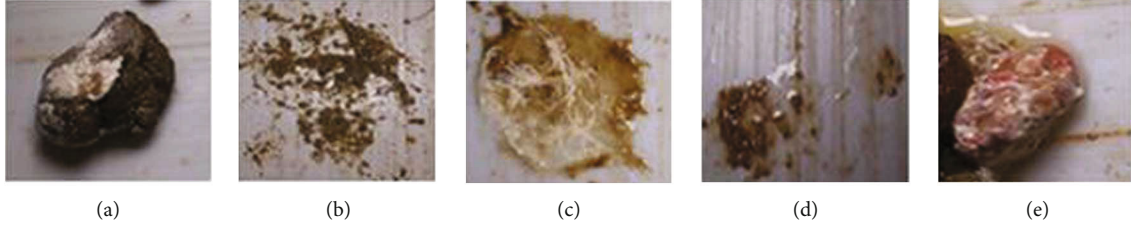


FIGURE 2: Data set labeling: (a) normal, (b) abnormal shape, (c) abnormal water content, (d) abnormal shape&water, and (e) abnormal color.

TABLE 2: The applied mage augmentation techniques.

Method	Description
Reversal	Flip the image horizontally or vertically at random
Blur	Gaussian blur (sigma between 0 and 3.0), average and or uniform blur (kernel size between 2 by 2 and 7 by 7), and median blur (kernel size between 3 by 3 and 11 by 11)
Sharpen	Sharpen each image and overlay the result with the original
Noise	Add Gaussian noise to some images
Gray-scale transformation	Add -10 to 10 to each pixel
Brightness	Change the brightness of images (50-150% of the original value)
Contrast normalization	Improve or worsen the contrast of images
Image rotation	Random rotation of $-15 \pm 15$ angles around the center of the image
Superpixel	Sample between 20 and 200 superpixels for each image and replace with the average value

where  $i$  is the anchor number and  $p_i$  is the probability of the  $i^{\text{th}}$  anchor number being the foreground. If the  $i^{\text{th}}$  anchor number is the foreground, then  $p_i, p_i^*$  is 1 or 0.  $t_i$  is the coordinate of the predicted bounding box.  $t_i, t_i^*$  is the coordinate of the ground truth.

In the ROI pooling, layer proposals generated by RPN are used to get a fixed size proposal feature map. Full connection operation can be used to identify and locate targets. There are two different sizes of proposals in the feature map, but after pooling, there are  $7 \text{ by } 7 = 49$  outputs, which can provide fixed-length input for the full connection layer.

**2.7. Faster R-CNN Training and Testing.** A four-step alternative training scheme was performed to train the model of Faster R-CNN.

- Step (1) Training the RPN. Region proposal task is performed by the RPN network, which is initialized and fine-tuned with an ImageNet pretrained model
- Step (2) Training the Fast R-CNN. The proposals generated by the initial step of RPN are classified by a separate detection network, which is also initialized by the ImageNet pretrained model

TABLE 3: The summary of the data set.

Label	Number of images		
	Training data set	Testing data set	Total
Normal	1100	277	1377
Abnormal shape	1463	366	1829
Abnormal color	1400	350	1750
Abnormal water content	1240	322	1562
Abnormal shape&water	895	224	1119
Total	6098	1539	7637

Step (3) Retraining the RPN. The detector network of the second step initialized the RPN network with fixed shared convolution layers. The RPN network is initialized with the detector network of step 2 with fixed shared convolution layers and only fine-tuned the unique layers of RPN

Step (4) Fine-tuning the Fast R-CNN. The unique layers of Fast R-CNN network are fine-tuned and generate precise bounding boxes of detected objects while the shared conv layers are kept fixed. In Step 1 and Step 2, two networks of RPN and Fast R-CNN share no parameter of the conv layers. However, in Step 3 and Step 4, the two networks share all the conv layers. For the whole process, Step 1 and Step 2 and Step 3 and Step 4 are adopted for alternative and continuous training

After the training of Faster R-CNN, the model was achieved; the detection of abnormal droppings in broilers only needs to use the final model generated by the training process and transmit the images forward.

**2.8. YOLO-V3 Architecture.** Compared with YOLO and YOLO-V2, the main improvements of YOLO-V3 are adjusting the network structure, using multiscale features to detect objects, and replacing the software Max with Logistic in object classification [17]. However, the YOLO-V3 [17] network was evolved from the YOLO [13] and YOLO-V2 [13] networks.

In image feature extraction, YOLO3 adopts a network structure called Darknet53 (which contains 53 convolution layers). In the Darknet53, from layer 0 to layer 74, there are 53 convolution layers and the rest are res layers (22); layers 75 to 107 are the YOLO layers. It draws on the residual



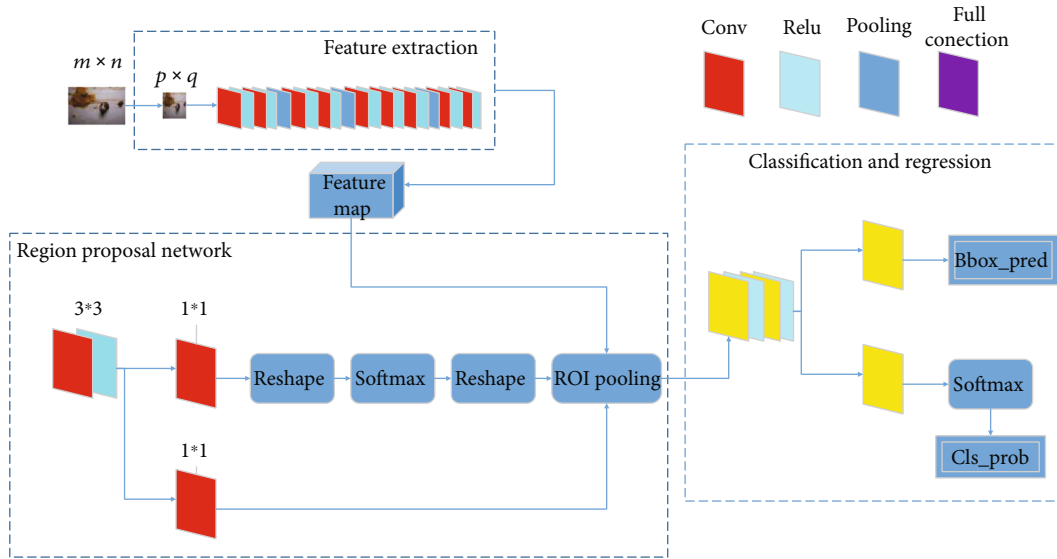


FIGURE 3: Overview of the proposed Faster R-CNN network structure.

network and sets up shortcut connections between some layers. Darknet53 consists of a series of 1 by 1 and 3 by 3 convolution layers (each convolution layer is followed by a BN layer and a Leaky ReLU). The YOLO-V3 feature interaction layer can be divided into three scales. The convolution network gets a scale detection result through several 32 times downsampling convolutions in Layer 79 [17]. Compared with the input image, the result is more suitable for detecting larger objects in the image because of the high downsampling multiple and the broader field of the feature map. For example, if the input is 416 by 416, the feature map will be 13 by 13. In order to achieve fine-grained detection, the feature map of layer 79 is upsampled and then fused with the feature map of layer 61. In this way, the fine-grained feature map of layer 91 is concatenated [17]. Similarly, after several convolution layers, the feature map of layer 79 is downsampled 16 times than that of the input image. It has a medium-scale sensing field and is suitable for detecting medium-scale objects. The feature map of layer 91 is upsampled and concatenated with the feature map of layer 36. Finally, the feature map is downsampled 8 times than that of the input image. It has the smallest receptive field and is suitable for detecting small objects [17].

The Res layer in the ResNet only does the shortcut operation of calculating the differences to determine if its input and output are consistent. To solve the problem of gradient dispersion or gradient explosion in the network, the Res layer proposes to change the layer-by-layer training of deep neural networks into step-by-step training [11]. It divides the deep neural network into several segments, each of which contains simple network layers. Then, the shortcut connection is used to train each segment for residual error and each segment learning part of the total loss. Finally, it achieves a small total loss. At the same time, the propagation of the gradient is well controlled to avoid the situation where the gradient disappears or explodes. The output layer directly outputs the recognition results (the type and location of the broiler drop-

pings by prediction). The output of the target box uses non-maximum suppression (NMS) to predict the filter box [12]. During the testing, the threshold of IOU is set and the only prediction box for each target with NMS is outputted. The IOU is defined by

$$\text{IOU} = \frac{\text{Overlap}}{\text{Union}}, \quad (3)$$

where Overlap is the intersection between the prediction box and the real box (label box) and the Union is the union of the prediction box and the real box. YOLO-V3 uses logistic output instead of the software Max to predict object categories to enable multiple tagging.

YOLO-V3 uses mean squared error as the loss function [13, 17]. It consists of three parts: coordinate error, IOU error, and classification error. It classified the droppings categories to detect abnormal droppings. The structure of the developed YOLO-V3 network is shown in Figure 4.

## 2.9. Optimization of YOLO-V3's Anchor Box

### 2.9.1. Anchor Box Selection Method.

All deep learning algorithms initialize the size and number of anchor boxes at an arbitrary value, for example, 9 in Faster R-CNN, 5 in YOLO-V2 [20], and 9 in YOLO-V3 [17]. However, the initialization of these anchor boxes is trivial because it affects model performance and computation resources. Therefore, it is imperative to select appropriate anchor boxes suitable for a particular data set. In this study, the *K*-means++ clustering algorithm was used to optimize anchor boxes of the YOLO-V3 algorithm for the data set of broiler droppings. The process of the optimization method is as follows. Firstly, the IOU of anchor boxes given by algorithm and each target box on each image in the training set was calculated. A point was randomly selected as the first initial point from the input set of training data points. For each point *X* in the training

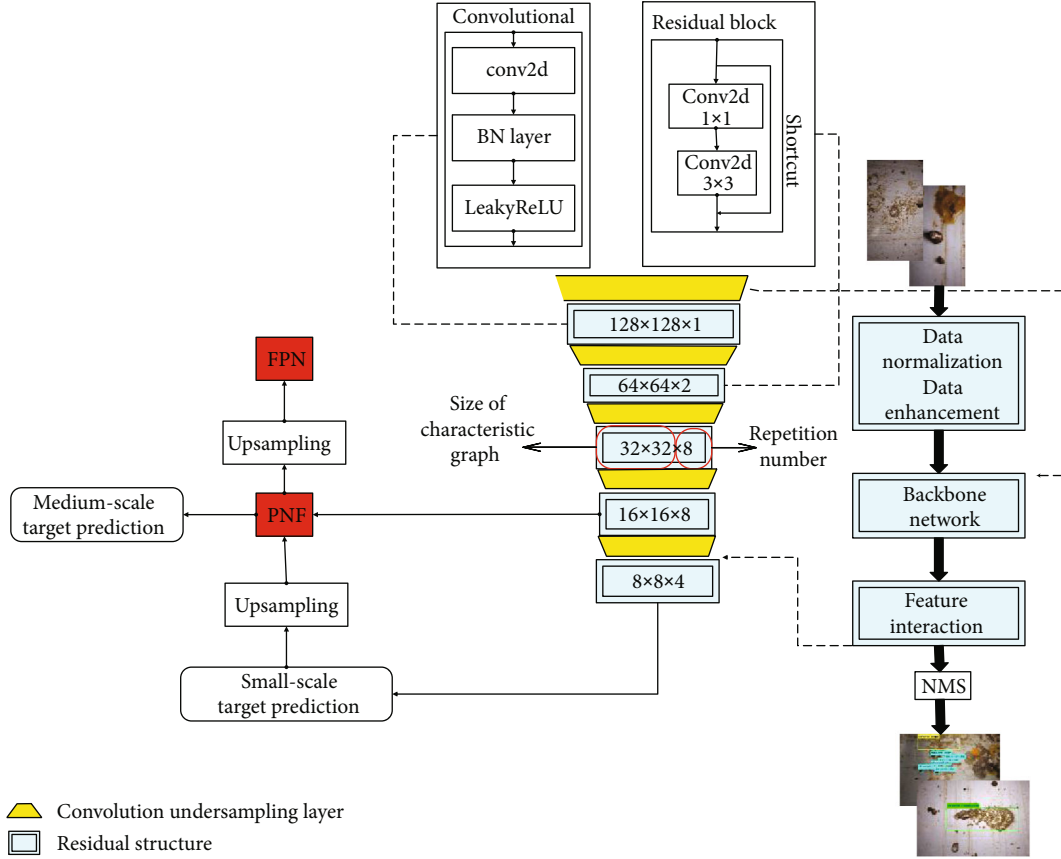


FIGURE 4: The overview of the proposed YOLO-V3 structure.

data set, the distance  $D(x)$  between the point  $X$  and the nearest clustering center was calculated. Selecting a new data point as a new clustering center, the rule to select was that the point with a larger  $D(x)$  was much more likely to be selected as the clustering center. The last two steps were repeated until all the clustering centers were selected. Then, for all points, the distance from these points to the clustering centers was calculated. If the point  $P_i$  is nearest to the cluster center point  $k_i$ , then  $P_i$  belongs to the group of  $K_i$  points. Next, move the clustering center to the center of its “point group.” Then, repeat the above steps until the clustering center point remains constant (moving distance being less than the set threshold). The flow chart of the algorithm is presented in Figure 5.

### 3. Results and Discussions

All the explored networks were trained by asynchronous Stochastic Gradient Descent (SGD). A distributed learning of the neural networks was attained by training in batch mode [21]. The batch number was set to 32 as the optimum level whereby a lower value such as 16 resulted in a much longer training time while a higher batch number resulted in high computational complexity. Similarly, the learning rate was set to 0.001; at a higher learning rate of 0.01, the network failed to converge, while at a lower learning rate of 0.0001, the network training needed more iterations to converge (longer training time).

**3.1. Faster R-CNN and YOLO-V3 Performance Evaluation.** The two detection models were designed and constructed based on the Tensorflow framework and the Darknet framework. They were trained and tested on a PC with NVIDIA GTX1080Ti GPU, 50 GB effective memory, and i7 7700k CPU with Ubuntu 16.04 operating system. To evaluate the performance of broiler droppings recognition models, recall, average precision (AP), and the mean value of average precision (mAP) were applied according to Equations ((4)), ((5)), and ((6)), respectively.

$$R = \frac{\text{TPos}}{\text{Pos}}, \quad (4)$$

$$\text{AP} = \int_0^1 P(R) dR, \quad (5)$$

$$\text{mAP} = \frac{\sum_{i=1}^c \text{AP}_i}{C}, \quad (6)$$

where TPos is the result of correct classification, Pos is the positive sample, and  $C$  is the number of sample classification. After 50,000 iterations of training, the models were tested on a testing data set. The testing results are presented in Table 4.

Table 4 indicates that YOLO-V3 had a longer training duration but with a much higher detection speed compared to Faster R-CNN. All the single class APs and recalls of Faster R-CNN were higher than that of YOLO-V3 at an average of

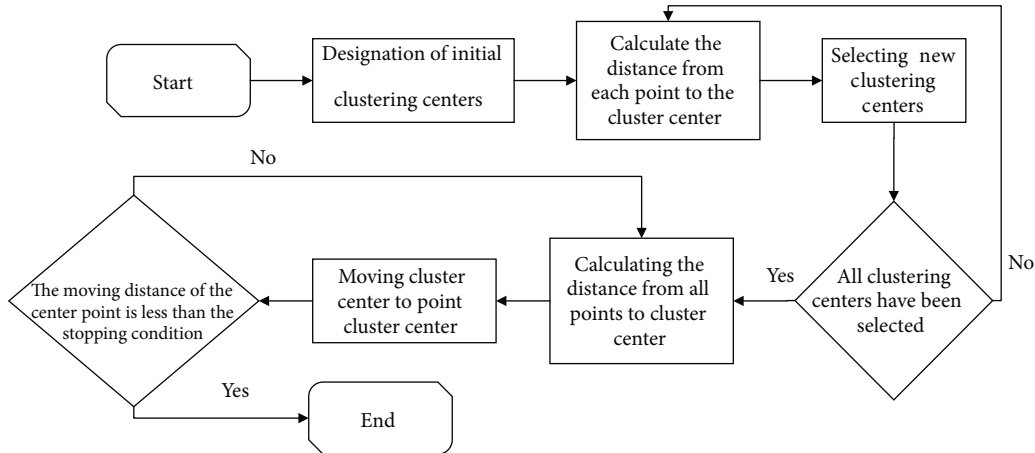


FIGURE 5: Flow chart of the anchor box selection algorithm.

TABLE 4: A comparison of the performance of YOLO-V3 and Faster R-CNN.

Recognition models	YOLO-V3		Faster R-CNN	
	Recall	AP	Recall	AP
Classes				
Normal	0.821	0.8285	0.989	0.8859
Abnormal shape	0.874	0.8813	0.996	0.9030
Abnormal color	0.922	0.8024	0.997	0.9815
Abnormal water content	0.927	0.8951	0.977	0.9822
Abnormal shape&water	0.895	0.8054	0.998	0.9125
Mean	0.887	0.8425	0.991	0.9332
Training duration (hr)	148		16	
Detection speed (fps)	43.5		5.7	

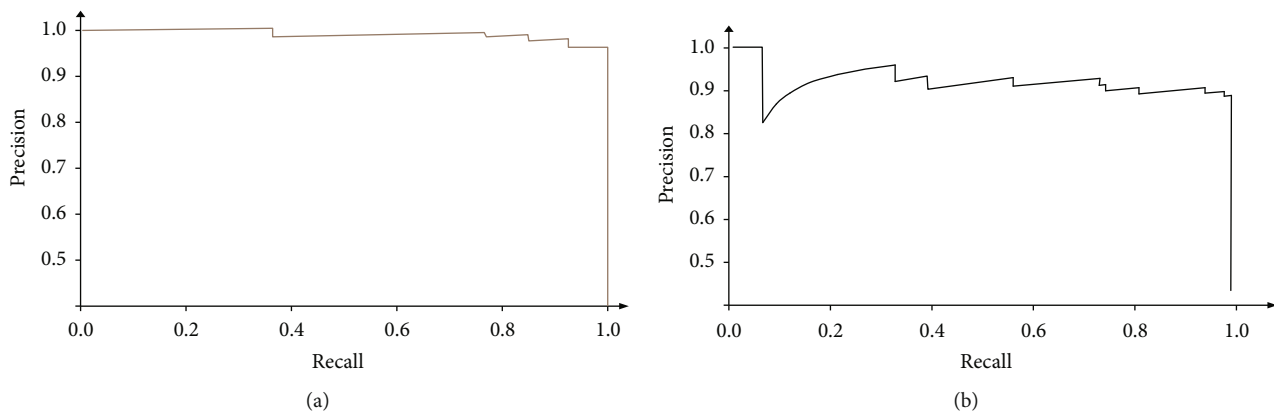


FIGURE 6: The precision-recall curves for (a) Faster R-CNN and (b) YOLO-V3.

0.9332 to 0.8425 and 0.991 to 0.887, respectively, on the whole testing set, considering that this study was performed on caged broilers with the droppings conveyor moving at a speed of  $0.09 \text{ ms}^{-1}$ . Hence, the frequency of image acquisition needed to be higher than 2 fps for both Faster R-CNN and YOLO-V3 systems to be applied.

The precision-recall curve for both models is presented in Figure 6. On comparison, Faster R-CNN performed better in recall and accuracy than YOLO-V3. In the low recall stage, it

was observed that the accuracy of YOLO-V3 was constant with an increase of recall. It then rises after a sharp decline; this indicated that the convergence on the testing set decreased rapidly after reaching the saturation point and the gradient disappeared or overfitting occurs in the process of backpropagation.

**3.2. Model Comparison at Different Iterations.** The depth of a network is critical to the performance of the Convolutional

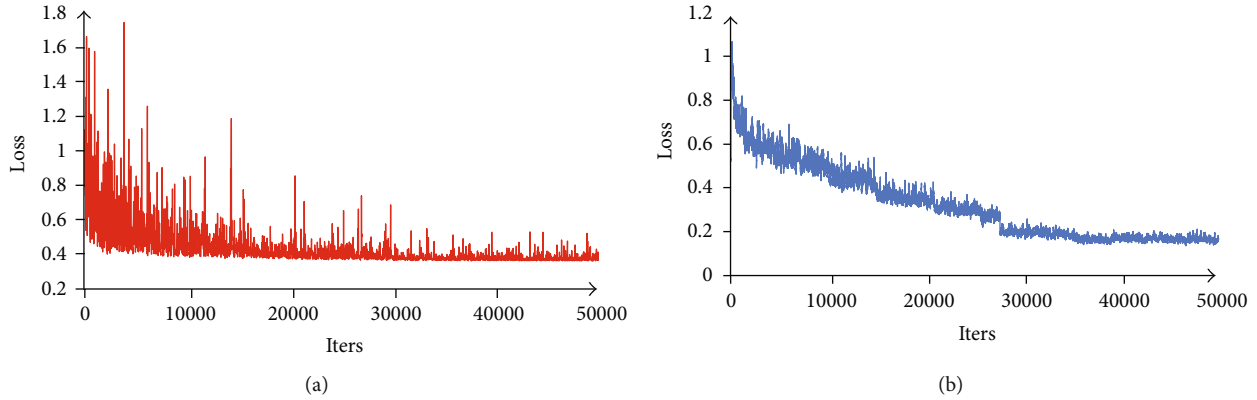


FIGURE 7: The average loss against the number of iterations for (a) Faster R-CNN and (b) YOLO-V3.

TABLE 5: Test results for 23,000 and 35,000 iterations of the YOLO-V3 weight model.

Iterations	Normal (ap)	Abnormal shape (AP)	Abnormal color (AP)	Abnormal water content (AP)	Abnormal shape&water (AP)	Map	Recall
23,000	0.780	0.859	0.799	0.860	0.764	0.813	0.846
35,000	0.837	0.863	0.806	0.871	0.778	0.831	0.879

Neural Network model, but with the increase of network layers, the overfitting problem arises [22]. Thus, the capacity of the data set, the depth of the convolution layer, and the number of iterations will affect the final model performance. In this study, the number of training data sets was 7637, the depth of YOLO-V3 was 106 layers except for the input layer, and Faster R-CNN applied ResNet101.

Figure 7 presents the average rate of loss value against the number of iterations for both networks. In the Faster R-CNN, the loss declined rapidly in the first 5000 iterations but with large fluctuations. The average loss value then stabilizes at 0.37 after about 30,000 iterations. However, for YOLO-V3, it is observed that the average loss decreased faster within the first 1000 iterations. At 20,000 to 25,000 iterations, the loss converged steadily at about 0.4. With increasing iterations from 25,000 to 27,000 iterations, there was a significant nonsmooth reduction in the loss. The loss is then observed to stabilize at 0.17 from 35,000 iterations. To evaluate the effect of the loss on recognition due to slow convergence, the weight models of YOLO-V3 in 23,000 iterations and 35,000 iterations were applied on the test data set. The test results are presented in Table 5.

It was established that the weighting model of 35,000 iterations was better than that of 23,000 iterations, whether in the accuracy of a single class or the overall map and recall. It can be concluded that (Figure 7(b) and Table 5) YOLO-V3 falls into the optimal local solution during the training process from 2000 to 25,000 iterations and achieved the optimal performance at about 31,000 iterations. Figure 8 presents the droppings classification based on Faster R-CNN during the model testing phase.

**3.3. Anchor Box Selection for YOLO-V3 Based on K-Means+.** Figure 9 presents the results of the optimization of YOLO-V3 with the K-means++ algorithm. Yellow stars were

the original anchor boxes, red triangles were the optimized anchor boxes, and blue points were the center point of training set GT boxes. The maximum number of iterations was set to 1000 and terminated if the clustering loss was less than  $1 \times 10^{-6}$  or the number of maximum iterations was reached. The average IOU of the anchor box and training set was 0.4292; the mean square deviation of IOU was 0.0819. Comparing with the initialized anchor box by YOLO-V3, IOU increased by 23.7%. With the same parameters and data set, YOLO-V3 was retrained using the optimized anchor box. Finally, the test results on the testing set and verification set were improved by 5% and 1.4%, respectively, by map 0.8846 and recall 0.899.

## 4. Conclusion

Digestive tract disease is one of the major diseases in broiler breeding. The monitoring technology of digestive tract diseases in broilers can promote the automation of the broiler breeding industry. Broiler droppings can be used as an index for evaluating digestive tract diseases in broilers. Broiler droppings were divided into normal and abnormal droppings, and abnormal droppings were divided into abnormal shape, abnormal color, abnormal water content, and abnormal shape&water. The performances of Faster R-CNN and YOLO-V3 on the broiler droppings data set were compared. Faster R-CNN achieved the best performance at about 30,000 iterations. On the testing set, the mAP was 93.32% and recall was 99.1%. YOLO-V3 achieved the best performance at about 31,000 iterations. On the testing data set, mAP was 84.25% and recall was 88.7%. By optimizing the anchor box of YOLO-V3 on the broiler droppings data set, the optimized anchor box's IOU was 23.7% higher than the initial anchor box, and the mean square deviation of the IOU was 0.085. Finally, Recall was increased by 1.4% and mAP is increased



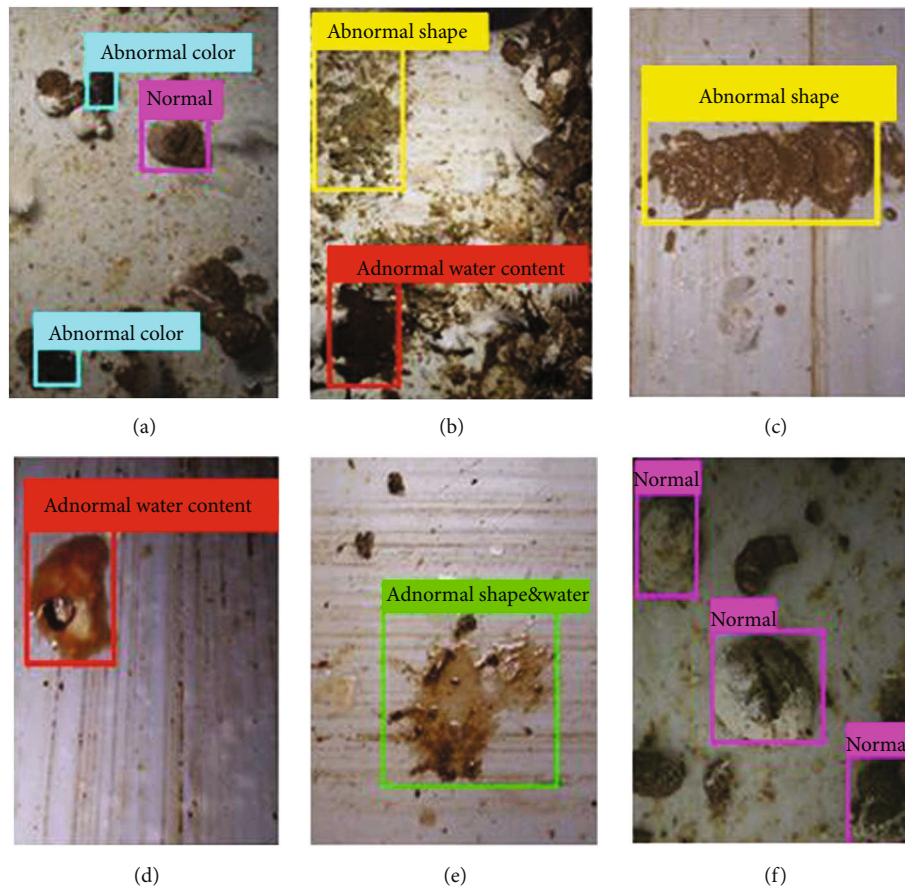


FIGURE 8: Broiler droppings label recognition during the testing phase based on Faster R-CNN (a) abnormal color and normal, (b) abnormal shape and abnormal water content, (c) abnormal shape and abnormal water content, (d) abnormal water content, (e) abnormal shape&water, and (e) normal.

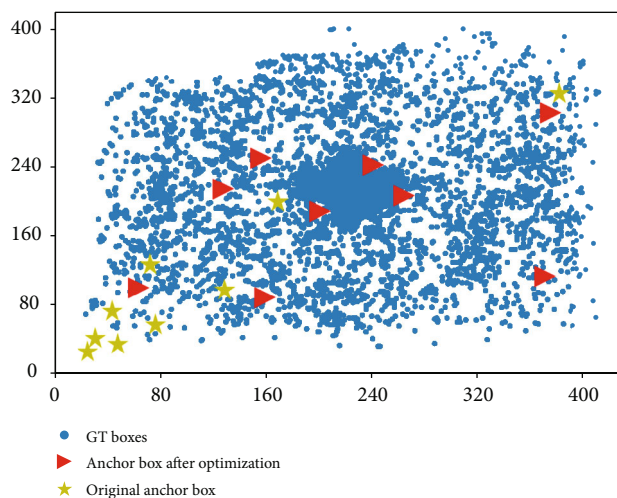


FIGURE 9: The scatter plot of the anchor box with and without optimization.

by 5%. This study provided an automatic and noncontact model for identifying and classifying abnormal droppings in broilers, which provided technical support for early warning of digestive tract diseases in broilers.

## Data Availability

The (image data set) data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

All the authors declare no conflict of interest.

## Authors' Contributions

All authors contributed equally to this paper.

## Acknowledgments

This research was funded by the National Key R&D Program of China (2017YFE0114400).

## References

- [1] T. Abbasi and S. A. Abbasi, "Reducing the global environmental impact of livestock production: the minilivestock option," *Journal of Cleaner Production*, vol. 112, Part 2, pp. 1754–1766.

- [2] P. Gerland, A. E. Raftery, H. ev ikova et al., “World population stabilization unlikely this century,” *Science*, vol. 346, no. 6206, pp. 234–237, 2014.
- [3] M. Pulido, J. Barrera-González, W. Badgery, J. Rodrigo-Comino, and A. Cerdà, “Sustainable grazing,” *Current Opinion in Environmental Science & Health*, vol. 5, pp. 42–46, 2018.
- [4] D. Berckmans, “Precision livestock farming technologies for welfare management in intensive livestock systems,” *Revue Scientifique et Technique*, vol. 33, no. 1, pp. 189–196, 2014.
- [5] M. Henchion, M. McCarthy, V. C. Resconi, and D. Troy, “Meat consumption: trends and quality matters,” *Meat Science*, vol. 98, no. 3, pp. 561–568, 2014.
- [6] G. Damerow, *The Chicken Health Handbook: A Complete Guide to Maximizing Flock Health and Dealing with Disease*, Storey Publishing, 2016.
- [7] X. Zhuang, M. Bi, J. Guo, S. Wu, and T. Zhang, “Development of an early warning algorithm to detect sick broilers,” *Computers and Electronics in Agriculture*, vol. 144, pp. 102–113, 2018.
- [8] X. Zhuang and T. Zhang, “Detection of sick broilers by digital image processing and deep learning,” *Biosystems Engineering*, vol. 179, pp. 106–116, 2019.
- [9] H. Shin, H. R. Roth, M. Gao et al., “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [12] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149.
- [13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: unified, real-time object detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, Las Vegas, NV, USA, June 2016.
- [14] S. Ren, K. He, R. Girshick, X. Zhang, and J. Sun, “Object detection networks on convolutional feature maps,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 7, pp. 1476–1481, 2017.
- [15] C. Zheng, X. Zhu, X. Yang, L. Wang, S. Tu, and Y. Xue, “Automatic recognition of lactating sow postures from depth images by deep learning detector,” *Computers and Electronics in Agriculture*, vol. 147, pp. 51–63, 2018.
- [16] Y. Tian, G. Yang, Z. Wang, H. Wang, E. Li, and Z. Liang, “Apple detection during different growth stages in orchards using the improved YOLO-V3 model,” *Computers and Electronics in Agriculture*, vol. 157, pp. 417–426, 2019.
- [17] J. Redmon and A. Farhadi, *YOLOv3: An Incremental Improvement*, 2018, <http://arxiv.org/abs/1804.02767>.
- [18] X. Yue, X. Hu, and L. Tang, “The influence of the amount of parameters in different layers on the performance of deep learning models,” *Computer Science and Application*, vol. 05, no. 12, pp. 445–453, 2015.
- [19] F. Wang, M. Jiang, C. Qian et al., “Residual attention network for image classification,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6450–6458, Honolulu, HI, United states, July 2017.
- [20] J. Redmon and A. Farhadi, “YOLO9000: better, faster, stronger,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6517–6525, Honolulu, HI, USA, July 2017.
- [21] L. Deng, D. Yu, and J. Platt, “Scalable stacking and learning for building deep architectures,” in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2133–2136, Kyoto, Japan, March 2012.
- [22] J. Schmidhuber, “Deep learning in neural networks: an overview,” *Neural Networks*, vol. 61, pp. 85–117, 2015.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

