IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 18, 2021, accepted May 14, 2021, date of publication May 17, 2021, date of current version May 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3081562

A New Few-Shot Learning Method of Digital PCR Image Detection

ZHANG BEINI[1], CHEN XUEE[1], LI BO[2], AND WEN WEIJIA[1]
[1]Department of Physics, Hong Kong University of Science and Technology (HKUST), Hong Kong
[2]Center for Soft and Living Matter, Institute for Basic Science (IBS), Ulsan 44919, South Korea

Corresponding author: Wen Weijia (phwen@ust.hk)

This work was supported in part by the SZSTI Shenzhen-Hong Kong Innovation Circle (Category D) 2019 under Grant SZST120SC15, and in part by the Zhongshan-Hong Kong University of Science and Technology Research Program under Grant ZSST20SC01.

ABSTRACT With the global pandemic of infectious diseases, the demand for accurate nucleic acid detection is daily increasing. The traditional threshold-based algorithms are adopted as the mainstream for processing the images of digital polymerase chain reaction (dPCR) now, but they are facing huge challenges on complex problems such as irregular noise, uneven illumination, and the lack of data. So, this paper proposed a novel few-shot learning method based on our *improved YOLOv3* model with fast processing speed and high accuracy to deal with complicated situations. Besides, to reduce the requirement of the large training dataset and annotation time of deep neural networks, we proposed the *Random Background Transfer Method* (RBTM) and *Source Traceability Annotation Method* (STAM) as the data augmentation and annotation method separately, which exploit the prior knowledge of the data and successfully realized the few-shot learning. Bases on the domain knowledge of dPCR images, our method could effectively augment images and reduce the labeling time by 70% while retaining the visually prominent features and improves the detection accuracy from 63.96% of the traditional threshold-based algorithm to as high as 98.98%. With the optimal processing speed and accuracy, our method is the state-of-art strategy for the detection of dPCR images now.

INDEX TERMS Digital PCR, few-shot learning, deep neural network, improved YOLOv3.

I. INTRODUCTION

With the global pandemic of the SARS-CoV-2-based disease (COVID-19), reverse transcription-polymerase chain reaction and real-time polymerase chain reaction, which adopt relative quantitative methods have exposed a serious shortcoming of the low accuracy. So, the demand for accurate disease detection is daily increasing [1], [2]. Actually, digital polymerase chain reaction (dPCR) adopts an absolute quantitative method and owns high accuracy which far beyond the relative quantitative polymerase chain reaction. According to the study of Valeria Cento, for patients who had false-negative results of the reverse transcription-polymerase chain reaction but with clinical symptoms, the detection results of the droplet dPCR which was verified by antibody detection have the accuracy of 100% and 95% for positive and negative results separately [3]. Therefore, dPCR is playing an increasingly important role in the detection of pathogens.

The associate editor coordinating the review of this manuscript and approving it for publication was Vishal Srivastava.

Academia and industry still adopt traditional threshold-based segmentation algorithms as the mainstream for dPCR at present [4]–[7]. These traditional algorithms have high requirements for image purity and simplicity, which is difficult to deal with complicated problems such as self-luminous irregular debris and uneven illumination [8]. When the brightness and size of the interfering targets are highly similar to the true positive targets, no matter how to adjust the settings of the threshold value, the accuracy is still far below the standard [9]. To reduce the problems of noise interference and uneven illumination caused by sample distribution and mechanical pressurization, the researchers put a lot of effort into improving the hardware performance [10], [11]. But this usually requires a higher cost of research and development, which makes the products more expensive. Indeed, these problems can be solved by more powerful algorithms. Some new algorithms such as the grid location and Mask Region Convolutional Neural Network (Mask R-CNN) have been proposed to improve the above weaknesses [8], [12]. But the grid location algorithm can only solve the noise problem of the chip-based dPCR. Mask R-CNN uses a two-stage method

of segmentation and detection which improves the accuracy at the expense of the processing speed, so it cannot adapt to the requirement of the high-throughput dPCR [13], [14]. However, with the global epidemic of many infectious diseases, high-throughput and real-time PCR will become more and more important for nucleic acid detection [15]–[18]. The You Only Look Once version 3 (YOLOv3) owns high detection accuracy and fast speed, which is the state-of-art algorithm for many areas such as fruit detection and ship detection [19], [20]. But this superior algorithm has not been brought into the field of dPCR.

Therefore, we proposed a novel method based on *improved YOLOv3*, *Random Background Transfer Method* (RBTM), and *Source Traceability Annotation Method* (STAM), which realized effective detection for the complex situation of the dPCR with relatively fast speed and high accuracy under the situation of the small training dataset. Most machine learning algorithms rely on classifiers for image recognition, while YOLOv3 uses a completely different approach. YOLOv3 is developed based on the darknet53 model and regards the classification problem as a single regression problem. Every single evaluation aims for the full image detection rather than proposed the specific ROI step by step. The simplification of the concept makes YOLOv3 have an extremely fast detection speed. Compared with Mask R-CNN, YOLOv3 could realize real-time object detection under the interference of complex noises [21], [22].

The method proposed in this paper establishes an improved YOLOv3 model based on a small set of experiment data, which realizes the state-of-art detection of the dPCR images at present. Compared with the other widely used methods from a comprehensive perspective, this method has the optimal accuracy, processing speed, and annotation time. It improves the accuracy of the complex situation from 63.96% of the traditional threshold-based algorithm to 98.98% with a false-positive rate of less than 1.5%. Compared with the Mask-RCNN, this method not only has comparable accuracy but also has 4.5 times faster's processing speed. Besides, the two newly proposed data augmentation and annotation methods in this paper, RBTM, and STAM, have improved the labeling efficiency by over 70% and solved a common dilemma of lacking high-quality data during the design stage of dPCR. Therefore, we could foresee the huge potential of this new few-shot learning method for the detection of dPCR images in the future.

II. RESULTS AND ASSESSMENT

A. DATASET RESOURCE AND PREPARATION

The droplet dPCR images with uneven illumination are acquired from the previous work of Wu *et al*. in Zhejiang University [23]. And the experimental dataset of Chip-based dPCR with complex irregular noise is derived from the work of our lab. Fluorescence reaction is produced by at least 35 thermal cycles of fluorescent droplet and luciferin respectively, then photographed by fluorescence microscope with 490nm green excitation light. The size and the number of reaction chambers of the chip-based dPCR are 100μm $\times$ 80μm (diameter $\times$ height) with 11200 chambers. After using the RBTM to augment the number of special images with complex noise and uneven illumination, 120 training images and 40 testing images of chip-based dPCR with the image resolution of 1600×1200 pixels, and 80 training images and 20 testing images of droplet dPCR with the image resolution of 1600×1200 pixels are finally selected.

The deep neural network requires a large number of samples to construct reliable models but the preparation of a large dPCR dataset is not only difficult but also time-consuming. Especially for the experiment of respiratory diseases such as the COVID-19, the experimenters need to wear protective suits and operate in P3 laboratories [24]. Weakly supervised deep learning algorithms based on scene-level tags have been proposed in recent years, which greatly reduce the time cost of data annotation and contribute to semantic segmentation [25], [26]. However, weakly supervised learning still requires a large number of training images, while the expense of acquiring digital PCR images is high and the detection of dPCR focuses on the instance segmentation rather than semantic segmentation. Few shot learning methods can achieve high accuracy with only a few training images, so it is more suitable for digital PCR. Therefore, we proposed a new data augmentation method, the Random Background Transfer Method in this paper to solve the common dilemma of lacking datasets for dPCR. Our few-shot learning method can complete learning by only 3 experimental images using a novel data augmentation method. Besides, a corresponding annotation method, Source Traceability Annotation Method, is developed to reduce the annotation time.

B. A NEW DATA AUGMENTATION METHOD

The Random Background Transfer Method proposed in this section realizes the effective data augmentation for the dPCR dataset. Generally speaking, deep neural networks need a large amount of data to ensure the quality of training [27], [28]. However, in practice, a large number of images are not produced in the stage of design but after the mass production of dPCR devices. Therefore, deep neural networks usually face the dilemma of lacking the dataset for newly designed tasks. If we want to exploit the advantages of deep neural networks to deal with complex problems for dPCR, an effective data augmentation method must be used.

The prior knowledge of the dPCR image makes it have the potential to realize few-shot learning. For the true-positive targets of dPCR, the distribution of different experiment results of the same layout design follows the Poisson distribution with similar brightness and shape. And there are often multiple true-positive targets in the same image, which could be utilized by the data augmentation method. However, the complex noise generated in the process of tablet pressing, centrifugation, and thermal cycling is not only different in shape but also unpredictable. So, the traditional data augmentation methods such as brightness adjustment, color balance transformation, and rotation have limited augmentation

efficiency and will change the visually prominent features of the true-positive targets, which are not suitable to realize few-shot learning for dPCR images [29]. Instead, we proposed a new data augmentation method, Random Background Transfer Method, in this section.

To ensure the accuracy of deep neural networks and reduce the cost of obtaining a large training dataset. The RBTM takes advantage of the highly regularized domain knowledge of the Chip-based dPCR to obtain the position and interval size as the parameters of the range function. Then use random.choice function to choose overwritten areas from the range. By adjusting the iteration times and overwritten size of the method, the different number of reaction chambers in the original image can be rewritten, thus changing the distribution of the original true-positive points while still preserving the visual features.

This method could select any pixel-wise RGB value of the original image as the selected value then apply it to rewrite other areas. The size and RGB value of the overwritten areas could be adjusted according to the actual demand optionally. As the examples of **Figure 1** show, if we select the value of the yellow box area in **Figure 1(a)** which represents the background, the more the iteration time increases, the less the true-positive targets are.

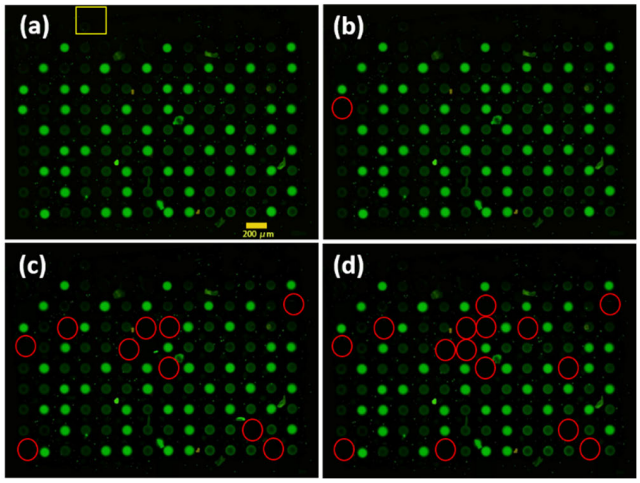

FIGURE 1. RBTM result. (a) Original experimental Image (b) Augmented image of 1 iteration. (c) 10 iterations. (d) 15 iterations.

The key parameters of RBTM could be calculated as the following formulas show:

$$X_{n_l} = X_1 + (n_l - 1)\frac{\sum_{i=2}^{m}(X_i - X_{i-1})}{(m-1)}$$
$$n_l \in \left\{1, 2, \cdots, \left[\frac{(L_a - X_1)}{l_b}\right]\right\} \quad (1)$$

$$Y_{n_w} = Y_1 + (n_w - 1)\frac{\sum_{i=2}^{n}(Y_i - Y_{i-1})}{(n-1)}$$
$$n_w \in \left\{1, 2, \cdots, \left[\frac{(W_a - Y_1)}{w_b}\right]\right\} \quad (2)$$

$$L_a = L - min\left(l_b, \frac{\sum_{i=2}^{n}(X_i - X_{i-1})}{(n-1)}\right) \quad (3)$$

$$W_a = W - min\left(w_b, \frac{\sum_{i=2}^{n}(Y_i - Y_{i-1})}{(n-1)}\right) \quad (4)$$

where L and W represent the length and width of the image, L_a and W_a represent the length and width of the max selective area. l_b and w_b represent the length and width of the transferred box. X_{n_l} and Y_{n_w} represent the top-left coordinate of transferred boxes. The m and n represent the target number of the first row and the first column separately.

$$f(i,j) = z(i,j)\,p(n_l, n_w) + u(i,j)\,(1 - p(n_l, n_w)) \quad (5)$$

$$p(n_l, n_w) = \delta\Big(0 < i - X_{n_l} < l_b \,\&\, 0 < j - Y_{n_w} < w_b \,\&\, u(i,j) - u\left(X_{n_l} + \frac{l_b}{2}, Y_{n_w} + \frac{w_b}{2}\right) < \varepsilon\Big) \quad (6)$$

where f is the augmented image, i is the x-axis of the image pixel and j is the y-axis of the image pixel, u is the original experimental image, z is the background image, δ equals 1 when the condition is satisfied and 0 otherwise. ε is the maximum allowable pixel value difference.

C. A NEW TIME-SAVING ANNOTATION METHOD

The Source Traceability Annotation Method proposed in this section successfully reduces the annotation time for deep neural networks by over 70%. After adopting the RBTM method, the newly generated images still have an intrinsic correlation with the original image, so we exploit this property to develop a new annotation method called Source Traceability Annotation Method.

The STAM requires that all images augmented from the same source should be stored in the same folder and named according to the number of iterations. In the duration of annotation, only the source needs to be marked fully manually. The rest of the augmented data produced by RBTM will select the annotation file which has the smallest absolute difference of the name to copy, and renames as the number of its iteration times, then reverse inputted the newly generated annotation file into the LabelImg or Labelme. Because the RBTM adopts linear iteration during the process of the background transfer, the chambers of dPCR Images will gradually decrease with the iteration times increasing. Due to the intrinsic relationship brought by the RBTM, for the augmented images, we only need to delete the redundant labels which belong to the covered chambers based on the newly generated annotation file. Therefore, the STAM could reduce the labeling time significantly.

As shown in **Figure 2,** the traditional annotation method which relies on fully manual annotation is time-consuming. After using STAM, the annotation time of Mask R-CNN and YOLO has been decreased by 75% and 70%, respectively. Although the labeling time of Mask R-CNN and YOLO was *both decreased prominently, the labeling time of YOLO is* shorter, which less than 35% of the Mask R-CNN.

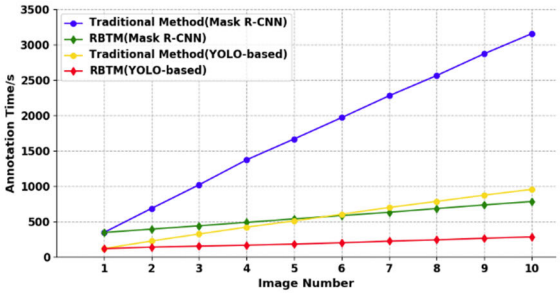

FIGURE 2. Comparison of the annotation time for chip-based dPCR.

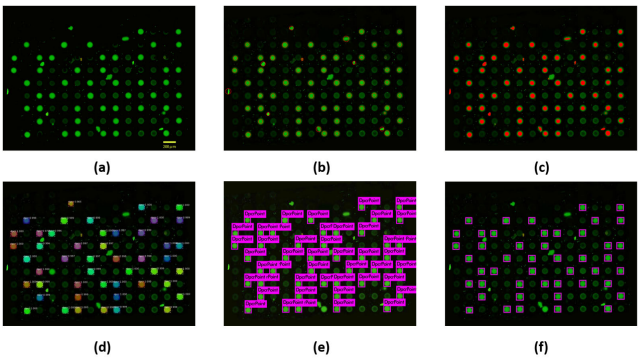

FIGURE 3. Detection results for irregular noise. (a) Image of Chip-based dPCR with irregular noise. (b) Comprehensive Threshold. (c) Simple Threshold. (d) Mask R-CNN. (e) YOLOv3 (f) Improved YOLOv3.

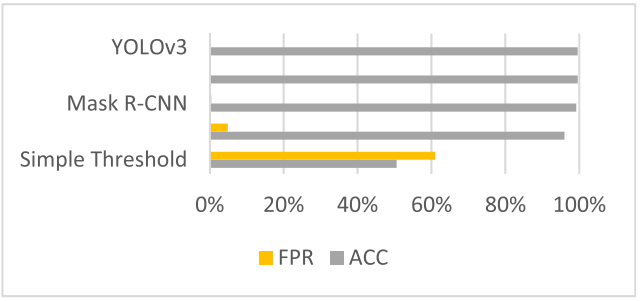

FIGURE 4. Accuracy comparison of Chip-based dPCR.

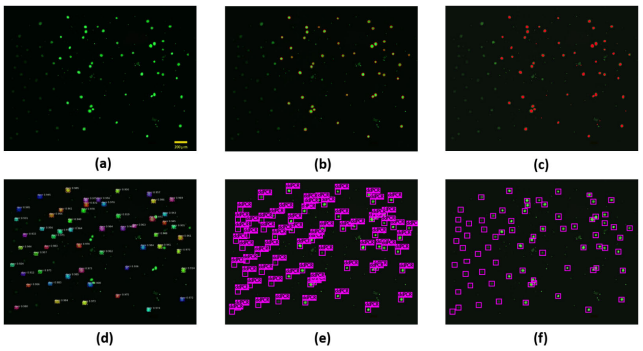

FIGURE 5. Detection results for uneven illumination. (a) Image of droplet dPCR with uneven illumination. (b) Comprehensive Threshold. (c) Simple Threshold. (d) Mask R-CNN. (e) YOLOv3 (f) Improved YOLOv3.

D. EVALUATION METRIC

Accuracy (ACC) and False Positive Rate (FPR) are one of the most important indicators of PCR detection. To compare the predictive performance of YOLO and threshold segmentation, the following formula is adopted for comparison in this paper [30]–[32]:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$FPR = \frac{FP}{FP + TN} \tag{8}$$

$$TPR = \frac{TP}{TP + FN} \tag{9}$$

For the predictive results of the algorithms, they can be divided into four categories: True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). TP means the positive target is correctly predicted to be positive, FP means the negative target is incorrectly predicted to be positive, FN means the positive target is incorrectly predicted to be negative, TN means the negative target is correctly predicted to be negative.

E. RESULTS OF INTERNAL IRREGULAR NOISE

The ideal dPCR image should be highly clean, but in actual experiments, the irregular noise is often produced after the process of centrifugation and thermal cycling, which greatly affects the recognition accuracy. As shown in **Figure 3**, the noise has complicated characteristics such as irregular, high-brightness, and similar size to the true positive targets.

For the Simple Threshold Segmentation algorithm, the accuracy is as low as 50.60% with the FPR as high as 61.08%. The Comprehensive Threshold Segmentation could successfully ignore the small noise, so it performs better than the Simple Threshold Segmentation. But if we focus on the 9 complicated noise, the Comprehensive Threshold Segmentation has the false-positive rate which exceeds 80%. On the contrary, the improved YOLOv3 and Mask R-CNN perform much better under the interference of the complicated noise, which owns the overall high accuracy of 99.60% and 99.20% with low FPR. Besides, our improved YOLOv3 model shows comparable accuracy and FPR with the YOLOv3.

F. RESULTS OF UNEVEN ILLUMINATION

Due to the complexity of the actual environment, the imbalance of illumination often occurs during the process of image preparation, which causes great trouble to target detection. As shown in **Figure 5**, the point light source of the image is placed on the upper right, resulting in the fluorescence intensity of the targets on the left are far below that of the right.

The threshold-based methods perform well on high-brightness targets but for the true positive targets affected by uneven illumination on the left, the accuracy of Simple Threshold segmentation and Comprehensive Threshold Segmentation was 0%. So, the overall accuracy of them is only 70% and 82.07%, separately. The Mask R-CNN performance better on the uneven area but the false-negative rate is higher than improved YOLOv3, so the overall accuracy is 86.90%.

TABLE 1. Performance comparison.

Assessment / Method	ACC	FPR	Processing Time/s(CPU)
Threshold	63.96%	41.97%	0.018
Comprehensive Threshold	71.64%	24.15%	0.372
YOLOv3	99.01%	1.69%	8.84
Mask R-CNN	93.20%	13.80%	23.55
Improved YOLOv3	98.98%	1.12%	5.25

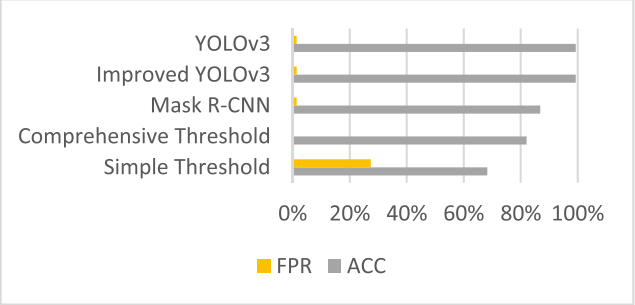

FIGURE 6. Accuracy comparison of droplet dPCR.

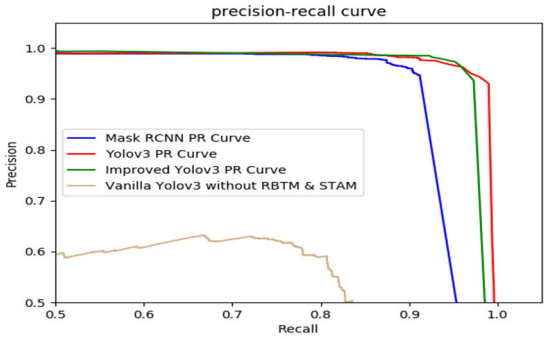

FIGURE 7. Precision-Recall curve comparison.

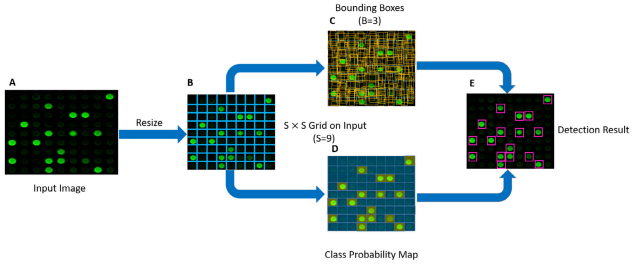

FIGURE 8. Workflow of improved YOLOv3.

While the improved YOLOv3 and YOLOv3 show the comparably optimal accuracy of 99.31% with FPR around 1.5%.

G. ASSESSMENT

The improved YOLOv3 owns the optimal detection accuracy and processing speed for complex situations. **Table 1** illustrates the average of accuracy, false-positive rate, and process time results of 40 validation images tested by CPU. From the aspect of the detection accuracy, Threshold Segmentation < Comprehensive Threshold Segmentation < Mask R-CNN < Improved YOLOv3 < YOLOv3. For processing speed, Mask R-CNN < YOLOv3 < Improved YOLOv3 < Comprehensive Threshold-Segmentation < Threshold Segmentation.

For the Precision-Recall (PR) curve for ablation study as **Figure 7** shows,Mask R-CNN, YOLOv3, and Improved YOLOv3 under the few-shot learning strategy own the Average Precision (AP) of 91.21%, 98.49%, and 97.61% respectively. The PR curve of vanilla YOLOv3 under the small dataset only has the AP around 54.68%, which far below the above methods and well illustrates the contribution of RBTM and STAM. From the aspect of the performance analysis, the curve of Mask R-CNN is completely wrapped by the curves of YOLO-based algorithms while the curve trends of Improved-YOLOv3 and YOLOv3 are very similar. So, the Yolo-based algorithms outperform the Mask R-CNN under the few-shot-learning strategy. And the Improved YOLOv3 has comparable performance with YOLOv3 while the computation speed is greatly improved.

Although the Threshold Segmentation has the fastest operation speed, it also has the worst detection accuracy for complex situations. The Comprehensive Threshold Segmentation could filter the noise which has a great difference in brightness or size with the targets, but the detection accuracy for complex noise and uneven illumination is still lower than deep learning methods. Mask R-CNN has higher accuracy for complex situations compares with traditional methods, but its computation speed is the slowest, which makes it difficult to carry out the real-time computation. The processing speed of our improved YOLOv3 method is around 1.68 times faster than YOLOv3 and around 4.5 times faster than Mask R-CNN. Besides, the annotation time of Mask R-CNN is 3 times longer than YOLO-based methods. And our method, improved YOLOv3, owns comparable accuracy with YOLOv3 and improves the processing speed by over 40%. In another word, when improved YOLOv3 uses the equivalent small dataset with other methods, it owns the optimal accuracy and processing speed, which has a huge potential to do real-time processing in the future.

III. METHOD AND STRUCTURE

A. METHOD OVERVIEW

We can understand the process of the detection as a regression problem. For the input image, the model will first resize the input image to the specific size then divide it as S × S grids, and each grid will produce B prediction results and confidence score. When the target is large and spans over multiple grids, the grid which the gravity center of the object falls into will be responsible for prediction. Although the prediction boxes are generated by a single grid, the reference information is from the global image. The improved YOLOv3 runs the convolutional network for the entire image and the predictions are produced based on the global image which could effectively avoid detecting the noise of the background as the target. For the example shown in the schematic diagram, each scale will produce 3 boxes so the totally tensor

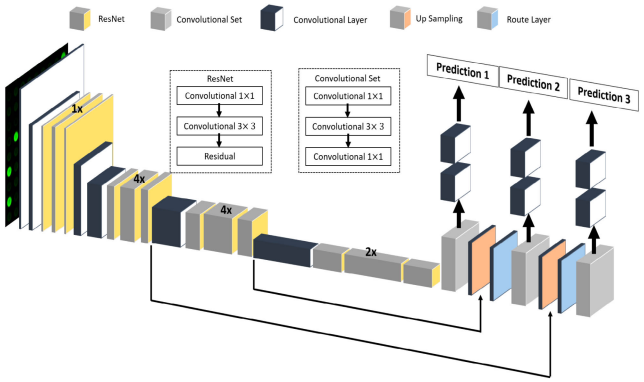

FIGURE 9. The structure of improved YOLOv3.

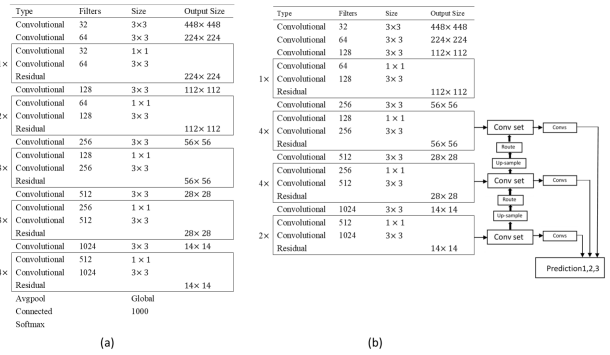

FIGURE 10. The backbone comparison. (a) Darknet53. (b) Improved YOLOV3.

number is $S \times S \times [3 * (4 + 1 + C)]$ and the C is the number of the class types [12]. Finally, the prediction results will be processed by model confidence and only the prediction boxes which exceed predetermined confidence will be output as the final detection results.

B. STRUCTURE AND PARAMETERS

The improved model proposed in this paper inherits the structure of a fully convolutional network from YOLOv3 and improved the residual networks and convolutional set as **Figure 9** shows to realize the balance between detection accuracy and processing speed. The improved YOLOv3 could be mainly divided into two types: Feature Pyramid Network (FPN) and prediction layer. The FPN is used for building and mining multi-dimensional information from images. And the prediction layer is used to effectively identify and judge the features mined by FPN. The predictions produced by the prediction layer complete the fusion of high-level features and low-level features through downward transmission, because the information of multiple dimension features could be contained in the same layer so it could reduce the judgment damage caused by single dimension information which completely separated with the features from upper layers.

The ordinary initial size for output size could be: 1024*1024, 640*640, 448*448, and 256*256 pixels. And the most commonly used one for YOLOv3 is 256*256 pixels.

TABLE 2. Key parameters.

Parameter	Value Setting
Channels	3
Momentum	0.9
Decay	0.0005
Learning Rate	0.0001
Max Batches	7000
Angle	0
Steps	5600,6300
Burn in	1000
Batch	64
Subdivisions	64
Width	448
Height	448
Filters in the Layer	18

However, this commonly initial size is too small for dPCR to do full image detection. This is because we utilize several down-sampling layers to improve processing speed, and most of our detection targets are small targets or medium targets that are less than 32*32 or between 32*32 and 96*96. If the initial value is too small, it will make the features of dPCR images cannot be extracted and recognized effectively. After several attempts, we found out that the best initial size of our dataset is 448*448.

Besides, YOLOv3 adopts logistic rather than softmax to do classification and the loss function for class prediction is the binary cross-entropy in (10), (11). The y represents the label with 1 for positive and 0 for negative:

$$loss = -\sum_{i=1}^{n} \hat{y}_i \log y_i + \left(1 - \hat{y}_i\right) \log\left(1 - \hat{y}_i\right) \quad (10)$$

$$\frac{\partial loss}{\partial y} = -\sum_{i=1}^{n} \frac{\hat{y}_i}{y_i} - \frac{1 - \hat{y}_i}{1 - y_i} \quad (11)$$

The softmax function, avgpool, and connected layers are designed for the Imagenet dataset particularly, so it is not necessary for our dataset. In order to avoid the loss of low-level features caused by pooling, instead of using any pool layers or connected layer, we adopted the fully convolutional network. For example, the function of down-sampling was realized by setting the stride of the convolutional layer to be 2.

The key parameters to train the model are shown in **Table-2**. Except for the requirement of small targets, specifying the length and width of the image could avoid any possible problems such as the slow processing speed caused by the failure of parallel GPU processing.

C. COMPARATIVE METHODS AND IMPLEMENTATION

In this paper, we adopt Simple Threshold Segmentation, Comprehensive Threshold Segmentation, and Mask R-CNN three methods as the comparative methods. Threshold Segmentation is the mainstream for processing PCR image in academia and industry at present, it only considers the targets whose brightness are higher than the threshold as valid and produce masks for them to measure later. On the basis of

the Threshold Segmentation, Comprehensive Threshold Segmentation brings in area selection to reduce the interference of regular fluorescent noise. It could successfully filter the noise which owns obviously different size or brightness with true positive targets but cannot deal with the noise which owns similar size and brightness with true positive targets. Mask R-CNN is a famous deep neural network that has been proposed in recent years, and it will first carry out a full image scan to propose the Region of Interest (RoI) that may contain the targets, then propose the classification and produce mask and detection boxes.

In our experiments, we adopt the function of adaptiveThreshold in OpenCV for Simple Threshold Segmentation. As for the Comprehensive Threshold Segmentation, we apply the function of Threshold and contourArea of OpenCV to calculate the threshold and produce masks. It could filter out the targets whose areas are much larger or smaller than the mean value. For Mask R-CNN, we utilize the open-source code from He Kaiming *et al.* [33], which needs Python 3.4, TensorFlow 1.3, and Keras 2.0.8 on Ubuntu 16.04 to process.

D. TRAINING AND ENVIRONMENT

Considering the applicability for portable PCR devices in the future, we utilized the deep learning workstation and common computers to process the training and testing separately. The training of deep neural networks requires high performances of computer graphic memory and workstation configuration. The hardware of the training computer is the PG620-P2G deep learning workstation with Intel Core I7-9800X processor and GeForce RTX 2080 Ti Graphic Card which possesses 11G graphic memory. The testing computer is the Precision 5820 which owns the processor of intel Xeon W-2245 and the whole procedure of testing is processed by CPU. The operating system is Ubuntu 16.04 and the project is written and tested on the darknet and anaconda3-4.4.0-Linux-x86 with python 3.6 and OpenCV2. The annotation tool labelImg needs a global operation, so it is installed under the Ubuntu main system disk with PyQt5. The vast majority of labelImg libraries can be installed uniformly as an integration package, but the libxcb-Xinerama library needs to be installed additionally to accommodate the Source Traceability Annotation Method's cross-environment tracking requirements.

We adopted the image dataset of red blood cells to pre-train the network. The training starts by entering the training command with the path of the config file, training dataset, and pre-trained weight. If it is necessary to visually analyze the loss and average IoU of the model, the corresponding command to save the training log shall be entered before the training. Saving the training log can check the source of problems when a large number of -Nan-regions appear and determine whether the model converges well, but it will slow down the training speed as well. So, it can be decided whether to save the training log according to the actual demand.

Yolo generates detecting results for all bounding boxes whose confidence level is above the threshold. By adjusting the threshold value, it can effectively select which detection boxes are visualized or not. Unless specially noted, 0.5 is adopted as the default threshold value in this paper, which means only the detection boxes with a confidence level greater than 0.5 will be considered as valid detection boxes and shown on the final detection result diagram.

IV. CONCLUSION

In this paper, we conducted the improved YOLOv3 model with the new data augmentation and annotation method for processing the dPCR images with irregular noise and uneven illumination under the situation of the small dataset. For this task, the average accuracy of the traditional threshold segmentation method is less than 65%, and the false positive rate is as high as over 40%, which makes it difficult to achieve effective detection. The disadvantages of deep neural networks compared to traditional algorithms are requiring more training samples and longer labeling time. But the RBTM and STAM proposed in this paper greatly reduce these disadvantages. It can effectively decrease the labeling time by more than 70% and rapidly augment the dataset without affecting the prominent visual features of dPCR. And the average accuracy of our improved YOLOv3 model achieves 98.98%, which shows obvious superiority among comparative methods. The processing speed of our model is 4.5 times faster and 1.68 times faster than Mask R-CNN and YOLOv3 separately, which realized the state-of-art for dPCR image detection. Besides, our model is more lightweight than Mask R-CNN and YOLOv3 so that it is more likely to be deployed on the embedded system and adapt to the requirements of high-throughput and real-time dPCR. In the future, we will work on using this model to achieve real-time development of dPCR on embedded and portable devices.

REFERENCES

[1] R. Lu, J. Wang, M. Li, Y. Wang, J. Dong, and W. Cai, "SARS-CoV-2 detection using digital PCR for COVID-19 diagnosis, treatment monitoring and criteria for discharge," *Med. Lett. CDC FDA*, p. 591, Apr. 2020.
[2] M. Ricchi, C. Bertasio, M. B. Boniotti, N. Vicari, S. Russo, M. Tilola, M. A. Bellotti, and B. Bertasi, "Comparison among the quantification of bacterial pathogens by qPCR, dPCR, and cultural methods," *Frontiers Microbiol.*, vol. 8, p. 1174, Jun. 2017.
[3] C. Alteri, "Detection and quantification of SARS-CoV-2 by droplet digital PCR in real-time PCR negative nasopharyngeal swabs from suspected COVID-19 patients," *PloS ONE*, vol. 15, no. 9, 2020, Art. no. e0236311.
[4] S. S. Reddi, S. F. Rudin, and H. R. Keshavan, "An optimal multiple threshold scheme for image segmentation," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-14, no. 4, pp. 661–665, Jul. 1984.
[5] A. C. Hatch, J. S. Fisher, A. R. Tovar, A. T. Hsieh, R. Lin, S. L. Pentoney, D. L. Yang, and A. P. Lee, "1-million droplet array with wide-field fluorescence imaging for digital PCR," *Lab Chip*, vol. 11, no. 22, pp. 3838–3845, 2011.
[6] X. Xu, S. Xu, L. Jin, and E. Song, "Characteristic analysis of Otsu threshold and its applications," *Pattern Recognit. Lett.*, vol. 32, no. 7, pp. 956–961, May 2011.
[7] J. Madic, A. Zocevic, V. Senlis, E. Fradet, B. Andre, S. Muller, R. Dangla, and M. E. Droniou, "Three-color crystal digital PCR," *Biomol. Detection Quantification* vol. 10, pp. 34–46, Dec. 2016.
[8] Z. Hu, W. Fang, T. Gou, W. Wu, J. Hu, S. Zhou, and Y. Mu, "A novel method based on a mask R-CNN model for processing dPCR images," *Anal. Methods*, vol. 11, no. 27, pp. 3410–3418, Jul. 2019.
[9] S. Rutsaert, K. Bosman, W. Trypsteen, M. Nijhuis, and L. Vandekerckhove, "Digital PCR as a tool to measure HIV persistence," *Retrovirology*, vol. 15, no. 1, pp. 1–8, Dec. 2018.

[10] S. H. Lee, J. Song, B. Cho, S. Hong, O. Hoxha, T. Kang, D. Kim, and L. P. Lee, "Bubble-free rapid microfluidic PCR," *Biosensors Bioelectron.*, vol. 126, pp. 725–733, Feb. 2019.
[11] J. M. Karlsson, T. Haraldsson, S. Laakso, A. Virtanen, M. Maki, G. Ronan, and W. van der Wijngaart, "PCR on a PDMS-based microchip with integrated bubble removal," in *Proc. 16th Int. Solid-State Sens., Actuator Microsyst. Conf.*, Jun. 2011, pp. 2215–2218.
[12] W. Jia, Y. Tian, R. Luo, Z. Zhang, J. Lian, and Y. Zheng, "Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot," *Comput. Electron. Agricult.*, vol. 172, May 2020, Art. no. 105380.
[13] K. He, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2961–2969.
[14] P. Bharati and A. Pramanik, "Deep learning techniques—R-CNN to mask R-CNN: A survey," in *Computational Intelligence in Pattern Recognition*. Singapore: Springer, 2020, pp. 657–668.
[15] J. Thornton, "COVID-19: Delays in getting tests are keeping doctors from work, health leaders warn," *BMJ*, vol. 370, Sep. 2020, Art. no. m3755.
[16] M. Xu, D. Wang, H. Wang, X. Zhang, T. Liang, J. Dai, M. Li, J. Zhang, K. Zhang, D. Xu, and X. Yu, "COVID-19 diagnostic testing: Technology perspective," *Clin. Transl. Med.*, vol. 10, no. 4, 2020, Art. no. e158.
[17] C. M. Hindson, J. R. Chevillet, H. A. Briggs, E. N. Gallichotte, I. K. Ruf, B. J. Hindson, R. L. Vessella, and M. Tewari, "Absolute quantification by droplet digital PCR versus analog real-time PCR," *Nature Methods*, vol. 10, no. 10, pp. 1003–1005, Oct. 2013.
[18] B. J. Hindson *et al.*, "High-throughput droplet digital PCR system for absolute quantitation of DNA copy number," *Anal. Chem.*, vol. 83, no. 22, pp. 8604–8610, 2011.
[19] 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," *Comput. Electron. Agricult.*, vol. 157, pp. 417–426, Feb. 2019.
[20] H. Cui, Y. Yang, M. Liu, T. Shi, and Q. Qi, "Ship detection: An improved YOLOv3 method," in *Proc. OCEANS, Marseille*, Jun. 2019, pp. 1–4.
[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788.
[22] E. Prasetyo, N. Suciati, and C. Fatichah, "A comparison of YOLO and mask R-CNN for segmenting head and tail of fish," in *Proc. 4th Int. Conf. Informat. Comput. Sci. (ICICoS)*, Nov. 2020, pp. 1–6.
[23] W. Wu, S. Zhou, J. Hu, G. Wang, X. Ding, T. Gou, J. Sun, T. Zhang, and Y. Mu, "A thermosetting oil for droplet-based real-time monitoring of digital PCR and cell culture," *Adv. Funct. Mater.*, vol. 28, no. 39, Sep. 2018, Art. no. 1803559.
[24] Y. Ichinose, S. Inoue, M. Shimada, G. Miring'u, B. Muriithi, A. Makumi, E. W. Apondi, M. B. Mwebia, C. Narita, S. Ashur, A. O. Kwallah, A. Galata, E. Odoyo, S. Huqa, M. Shah, M. Karama, and M. Horio, "Operations at biosafety level III: The P3 laboratory," in *Epidemiology II: Theory, Research and Practice*. Hong Kong: iConcept Press, 2015.
[25] Y. Li, W. Chen, Y. Zhang, C. Tao, R. Xiao, and Y. Tan, "Accurate cloud detection in high-resolution remote sensing imagery by weakly supervised deep learning," *Remote Sens. Environ.*, vol. 250, Dec. 2020, Art. no. 112045.
[26] Y. Li, Y. Zhang, X. Huang, and A. L. Yuille, "Deep networks under scene-level supervision for multi-class geospatial object detection from remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 146, pp. 182–196, Dec. 2018.
[27] B. Ma, X. Wei, C. Liu, X. Ban, H. Huang, H. Wang, W. Xue, S. Wu, M. Gao, Q. Shen, M. Mukeshimana, A. O. Abuassba, H. Shen, and Y. Su, "Data augmentation in microscopic images for material data mining," *Npj Comput. Mater.*, vol. 6, no. 1, pp. 1–9, Dec. 2020.
[28] Y. Fujisawa, Y. Otomo, Y. Ogata, Y. Nakamura, R. Fujita, Y. Ishitsuka, R. Watanabe, N. Okiyama, K. Ohara, and M. Fujimoto, "Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis," *Brit. J. Dermatol.*, vol. 180, no. 2, pp. 373–381, Feb. 2019.
[29] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Goncalves, W. R. Schwartz, and D. Menotti, "A robust real-time automatic license plate recognition based on the YOLO detector," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–10.
[30] Y. Wang, C. Barbacioru, F. Hyland, W. Xiao, K. L. Hunkapiller, J. Blake, F. Chan, C. Gonzalez, L. Zhang, and R. R. Samaha, "Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays," *BMC Genomics*, vol. 7, no. 1, pp. 1–16, Dec. 2006.
[31] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," in *Proc. NESUG Health Care Life Sci.*, Baltimore, MD, USA, vol. 19, 2010, p. 67.
[32] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.
[33] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*. [Online]. Available: http://arxiv.org/abs/1804.02767

ZHANG BEINI received the B.E. degree in electronic information engineering from Chongqing University, Chongqing, China, in 2018, and the M.Sc. degree in electronic engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2019, where she is currently pursuing the Ph.D. degree.

Her main research interest includes utilizing deep neural networks to deal with the complex problems of microscopic images.

CHEN XUEE received the B.E. and M.E. degrees from Chongqing University, Chongqing, China, in 2014 and 2017, respectively. She is currently pursuing the Ph.D. degree with the Hong Kong University of Science and Technology.

Her main research interests include micro/nanofluidics and biosensors.

LI BO received the B.E. degree from Wuhan University, China, in 2012, and the Ph.D. degree from the Hong Kong University of Science and Technology, in 2017.

Since 2017, he has been a Postdoctoral Research Fellow with the Center for Soft and Living Matter, Institute for Basic Science, South Korea. He has published 11 articles in his research fields, including soft condensed matter, biophysics, glass physics, and imaging processing. He received the Young Scientist Award of the Hong Kong University of Science and Technology during 2015–2016 academic year.

WEN WEIJIA received the B.E. and M.E. degrees from Chongqing University, Chongqing, China, in 1992 and 1995, respectively, and the Ph.D. degree from the Institute of Physics, Chinese Academy of Sciences, China, in 1995.

From 1995 to 1999, he was a Research Fellow with the Hong Kong University of Science and Technology (HKUST), Hong Kong, and UCLA. Since 1999, he has been an Academic Staff Member with the Department of Physics, HKUST. He has published more than 400 articles in his major research fields, including more than 350 journal articles and 50 conference invited report. He has been published more than 300 SCI articles. His refereed journal articles have been cited more than 10,500 times resulting in an H-index of 52. His main research interests include soft condensed matter physics, electrorheological (ER) and magnetorheological (MR) fluids, field-induced pattern and structure transitions, micro- and nano-fluidic controlling, advanced functional materials, including microsphere and nanoparticle fabrications, thin film physics, band gap materials, metamaterials, and nonlinear optical materials.