IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 29, 2019, accepted August 9, 2019, date of publication August 15, 2019, date of current version September 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2935613

Mobile Communication Base Station Antenna Measurement Using Unmanned Aerial Vehicle

YIKUI ZHAI[1,2], (Member, IEEE), QIRUI KE[1], YING XU[1], WENBO DENG[1], JUNYING GAN[1], JUNYING ZENG[1], WENLVE ZHOU[1], FABIO SCOTTI[2], (Member, IEEE), RUGGERO DONIDA LABATI[2], (Member, IEEE), AND VINCENZO PIURI[2], (Fellow, IEEE)

[1]Department of Intelligent Manufacturing, Wuyi University, Jiangmen 529020, China
[2]Departimento di Information, Universitá degli Studi di Milano, 20133 Milano, Italy

Corresponding author: Ying Xu (xuying117@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771347, in part by the Characteristic Innovation Project of Guangdong Province under Grant 2017KTSCX181, in part by the Young Innovation Talents Project of Guangdong Province under Grant 2017KQNCX206, in part by the Jiangmen Science and Technology Project under Grant [2017] 268, in part by the Youth Foundation of Wuyi University under Grant 2015zk11, in part by the Opening Project of GuangDong Province Key Laboratory of Information Security Technology under Grant 2017B030314131, in part by the 2018 Opening Project of Guangdong Province Key Laboratory of Digital Signal and Image Processing, and in part by the Basic Research and Applied Basic Research Key Project in general colleges and Universities of Guangdong Province under Grant 2018KZDXM073.

ABSTRACT Traditional base station antenna measurement methods conducted with professional worker climbing towers tend to raise safety and inefficiency concerns in practical application. Designed to address the above problems, this paper proposes an intelligent and fully automatic antenna measurement unmanned aerial vehicle (UAV) system for mobile communication base station. Firstly, an antenna database, containing 19,715 images, named UAV-Antenna is constructed by image capturing with the help of UAVs flying around various base stations. Secondly, Mask R-CNN is adopted to train an optimal instance segmentation model on UAV-Antenna. Then, pixel coordinates and threshold are utilized for measuring antenna quantity and separate all antenna data for further measuring. Finally, a least squares method is employed for measuring antenna parameters. Experimental results show that the proposed method can not only satisfy the industry application standards, but also guarantee safety of labors and efficiency of performance.

INDEX TERMS UAV, antenna measurement, instance segmentation, least squares.

I. INTRODUCTION

Mobile communication base station serves to transmit radio transmission and reception stations between mobile communication switching center and mobile terminal within a certain radio coverage area. Down-tilt angle of antenna of mobile communication base station is set according to the coverage requirements of the network, which correlates with topography, size of traffic and quality of network service. To meet public's communication needs, an increasing number of base stations are built, leading to an exponential growth in the need of antenna down-tilt angle adjustments. Traditional antenna measurements are measured by professional workers with rulers, yet, due to the massive number of base stations, the high possibility of deviation of manual measured data and the complex environment, it is extremely difficult to guarantee the timeliness of data and the safety of professional workers.

Thus, out of the concerns mentioned above, traditional antenna measurement method gradually loses people's preference. With growing attention on safety issue, some researches have proposed multiple methods to measure antennas [1], [2]. Geise *et al.* [1] introduced a portable near-field antenna measurement system that could accurately measure the position and orientation of the probe antenna during the near-field scan by 6D laser tracking system. Garcia-Fernandez *et al.* [2] developed an unmanned aerial system for antenna measurements that used a real-time dynamics meter and a laser altimeter to geographic matching the measurements on centimeter level. What mentioned above are state-of-the-art methods in antenna parameter measure-ment. However, they all share the same features as strict hardware requirements and high capital consumption, which makes them difficult to apply in practice.

The associate editor coordinating the review of this article and approving it for publication was Dongxiao Yu.

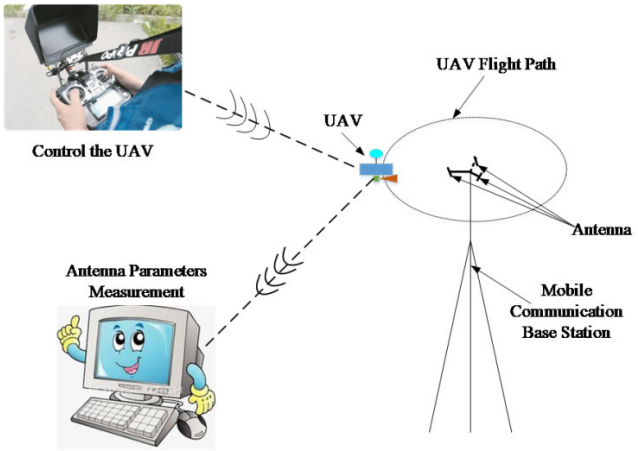


FIGURE 1. Antenna parameters measurement system using UAV.

This paper presented an automatic antenna parameters measurement method based on deep learning, which consumes little, engages with low hardware requirements and suits for popularization and its performance surpasses all other state-of-the-art methods. The flow chart of proposed method can be referred to Figure 1.

Deep learning method was proposed by Hinton in 2006, and has gradually received attention in massive information processing, image feature extraction and prediction modeling [3]–[7]. It is essentially an unsupervised layer-by-layer training method, which uses unlabeled samples for pre-learning, corrects and learns the discriminated features through a small number of labeled samples that has achieved amazing performance in object detection, segmentation and recognition.

What surprised us is that Mask R-CNN proposed by He *et al.* [32] combined object detection and semantic segmentation, and suggested efficient instance segmentation. Inspired by the above advancement, this paper presented a mobile communication base station antenna measuring method by using UAV. To begin with, an antenna database named UAV-Antenna is established with 19,715 communication base station images. Secondly, Mask R-CNN is adopted to train the training set of UAV-Antenna to obtain an optimal instance segmentation model. The testing results of this optimal model are shown in Figure 2(better viewed in color), where (a)∼(c) are the original antenna images and (d)∼(f) are the segmented antenna images. Then, pixel coordinates and threshold are utilized for measuring antenna quantity and separate all antenna data for further measuring. Finally, a least squares method is employed for measuring antenna parameters. Our major contributions can be summarized as follows:

(i). A method named mobile communication base station antenna measurement using UAVs has been proposed that is faster and safer than the previous state-of-the-art and can address the problems of inefficiency and randomness in traditional methods.

(ii). The core of proposed method combines Mask R-CNN, Least Square, frame sequence analysis with UAV to achieve pixel-level antenna parameters automatic measurement of mobile communication base station.

(iii). To meet industry standards and achieve high detection accuracy and fitting accuracy, an antenna database containing 19715 images named UAV-Antenna has been constructed by image capturing with UAVs flying around various base stations. These images consist of a training set of 19496 unlabeled images and a testing set of 219 labeled images.

(iv). Experiments include time, detection and fitting accuracy analysis on models with varying parameters setting, and UAV-Antenna database is tested with several architectures in terms of YOLOv3, Faster R-CNN and Mask R-CNN.

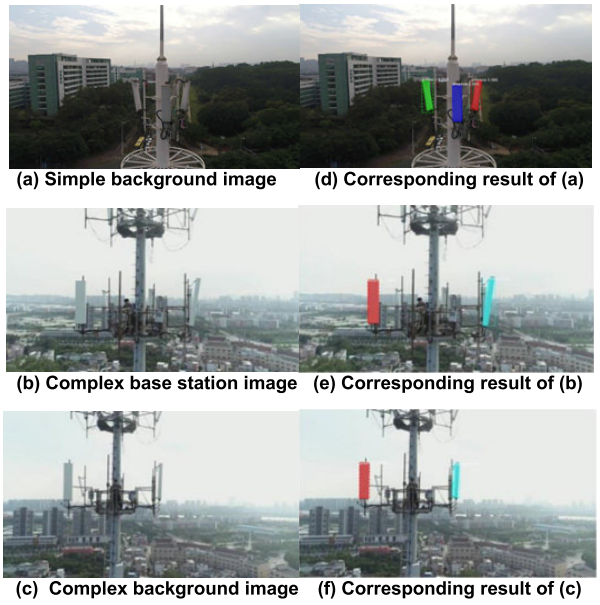


FIGURE 2. (a)∼(c) are original antenna images and (d)∼(f) are segmented antenna images.

II. RELATED WORKS

UAVs are unmanned aerial vehicles operated by radio remote control equipment and self-contained program control devices. In recent years, industries in mounting number begin to utilize UAVs instead of traditional labors. For instance, agriculture uses UAVs for fertilization, industry for inspections, and the film and television industry for aerial photography, special effects and so forth. Shakhatreh *et al.* [8] introduced the capability of using small low cost UAV system to measure and diagnose field antennas. Shahzad *et al.* [9] conducted a survey on the civilian application of UAV and the challenges that human is facing. Nonetheless, in the measurement of mobile communication base station antennas, the use of only UAVs may immensely increase the hardware cost and the flying time of UAVs may be greatly shortened with their ascending weight. Therefore, introducing machine vision with the assistance of UAV to measure antenna parameters becomes the optimal solution.

Object detection is one of the core issues in the field of computer vision that has received extensive attention in many fields such as industry, agriculture, and manufacturing [10]–[14]. Girshick [15] certified a region-based fast convolutional network object detection method, with the use of a deep convolution network to effectively classify object candidate regions. Liu *et al.* [16] validated a method for detecting objects in an image using a single deep neural network, which discretizes the output space of bounding box into a set of default boxes, each of which has different scales and proportions. When forecasting, the network generates a score for each object category in each default box and adjusts the box to better match object shape. Redmon *et al.* [17] originated a method of object detection, shifting the perspective from the problem of object detection to the problem of spatially separated bounding box and related class probability regression. Ren *et al.* [18] introduced a more advanced object detection network that relies on a candidate region algorithm to assume an object location. Region Proposal Network (RPN) also is presented, which shares a full image convolution features with the detection network to generate nearly free candidate regions. However, object detection can only detect antenna but cannot measure antenna parameters. Consequently, it is hoped that the relationship between pixel and antenna parameters can be favorably identified via introducing semantic segmentation.

Semantic segmentation is a pixel-level classification of the target images. It has been widely applied in the fields of geographic information systems, unmanned vehicle driving, medical image analysis, and robotics [19]–[23]. Chen *et al.* [24] introduced a model for improving segmentation results by adding a simple and efficient decoder module that applies depth separable convolution to spatial pyramid pools and decoder modules so as to form faster and stronger encoder-decoder network. Wang *et al.* [25] raised a method to improve pixel-level semantic segmentation by manipulating convolution-related operations. This method creates a dense up-sampling convolution to generate pixel-level predictions, which captures and decodes more detailed information that is usually omitted in bilinear up-sampling. Moreover, a hybrid expansion convolutional framework is proposed at the coding stage. The receiving domain of the network is expanded to aggregate global information. Chen *et al.* [26] suggested the use of convolution of upsampling filters to achieve dense prediction tasks, and a shrinking spatial pyramid pool for robust segmentation of objects on multiple scales. Long *et al.* [27] proposed a method to establish a full convolutional network, which inputs objects of any size and outputs a correspondingly sized outcome by effective reasoning learning. Furthermore, it transforms the widely used classification network (AlexNet, VGG net, and GoogLeNet) into a full convolutional network, and continues defining a mobile framework that will combines semantic information from deep layers with appearance information from shallow layers to produce accurate and detailed segmentation. After semantic segmentation, the relationship between antenna parameters and pixel points can be easily acknowledged. As for the specific parameter value of antenna, linear fitting method is utilized to quantize antenna parameters which can readily address this problem.

Linear fitting refers to the functional relationship between a continuous curve and a coordinate represented by a discrete set of points on a same plane. More broadly, the corresponding problem in space or high-dimensional space also falls into this category. In the process of numerical analysis, linear fitting is the use of analytical expressions to fit discrete data points. Linear fitting, as a common method in mathematical calculations, has been applied in architecture, physics, chemistry, and even artificial intelligence [28].

By now, no predecessor has ever combined object detection, semantic segmentation and linear fitting in antenna parameter measurement. For the first time, this paper validated a fully automatic antenna parameter measurement method based on instance segmentation [29-31], least squares, frame sequence analysis and UAV [32-35], which enjoys remarkable preciseness, rapid recognition and outstanding performance.

III. UAV-ANTENNA DATABASE

An antenna database named UAV-Antenna was established in order to improve the instance segmentation accuracy of mobile communication base station antenna and pixel-level antenna parameters quantization accuracy in different environments, for instance, low light situation, complex background and so forth. This antenna database consists of 19,715 images, captured with UAVs flying around various base stations. These images contain a training set of 19,496 images and a testing set of 219 images. Among them, the labels of testing set were manually adjusted by professional workers to be compared with the results obtained via the methods proposed in this paper.

A. DATA CAPTURING

In the data capturing phase, two UAVs were used to fly around the mobile communication base station with various background complexity in different time periods, for instance morning, noon and evening, and obtained video data of base station through HD cameras provided on UAVs. The flight parameters of UAVs are strictly controlled as follows:

(1) When UAV flying around the mobile communication base station, its flying height must be the same as the height of center position of antenna of communication base station to avoid inaccurate pixel-level quantization of antenna parameters due to the inconsistency of view.
(2) The UAVs' flying radius was adjusted as 5 to 6 meters to ensure the appearance of the base station in the field of view and the visibility of antenna outline to improve the detection accuracy of the proposed method.
(3) The angular velocity of the UAVs stands at 3 degree/s, ensuring every frame in the video clear enough to be detected.

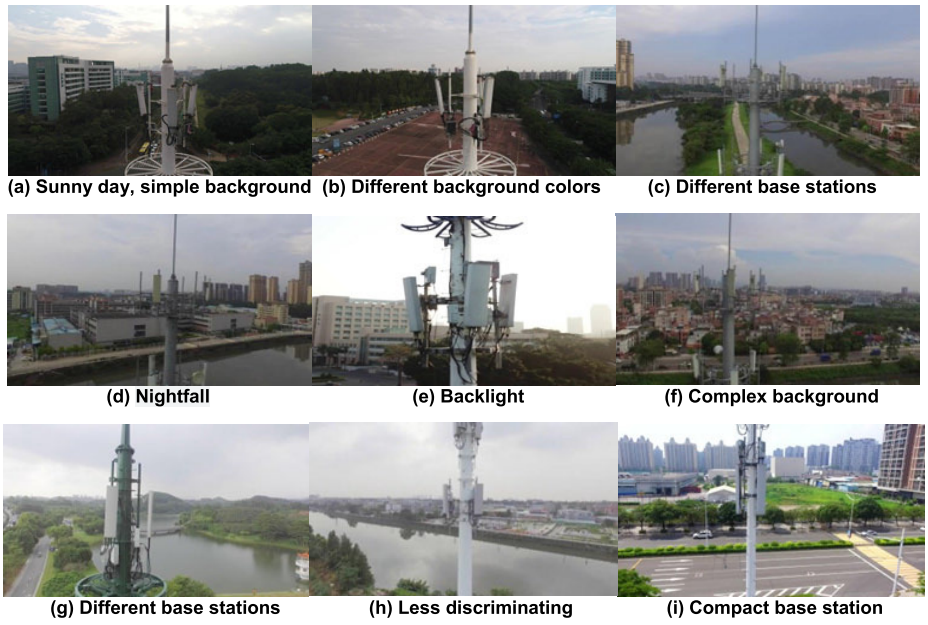


FIGURE 3. (a)~(i) are samples of UAV-antenna on various lighting and backgrounds.

B. DATA PROCESSING

In the data processing phase, with the video data of communication base station collected during the data capturing phase, interval frame sampling process was conducted and ensured that each image obtained from the data capturing phase has at least one antenna that can be detected. The database consists of images from diverse base stations with various environments and background complexity to ensure the accuracy and universality of the optimal model. It turns out that 19,496 antenna images were sampled to form the training set of UAV-Antenna. Parts of training set images are shown in Figure 3(better viewed in color).

C. DATA LABELING

Training set images were marked in data labeling stage. Specific steps are presented as follows: First and foremost, VGG Image Annotator was used to mark antenna outline on the front, side and front side of the antennas in each image, with name tag read ''antenna''. After all of the 19496 images had been marked, a json file suitable for training of Mask R-CNN was generated and placed in the training set to have the same path as training images.

IV. PROPOSED AUTOMATIC ANTENNA PARAMETERS MEASURING METHOD

Aiming to solve the drawbacks of traditional manual measurement of mobile communication base station antenna parameters, for instance low efficiency, high mortality and other measurement difficulties, this paper pioneered with an intelligent and fully automatic antenna measurement method using UAV system for mobile communication base station, which unites Mask R-CNN, linear fitting, frame sequence analysis and UAV to realize antenna parameters pixel-level measurement.

The following part proves to specify this method: above all, an antenna database named UAV-Antenna containing 19,715 communication base station images is constructed by image capturing with UAVs flying around various base stations. Secondly, Mask R-CNN is adopted to train an optimal instance segmentation model on UAV-Antenna database. Then, pixel coordinates and threshold are utilized for measuring antenna quantity and separate all antenna data for further measuring. Eventually, least squares method is employed for measuring antenna parameters. Experimental results show that the proposed method outperforms methods of measuring antenna parameters with practical hardware. Additionally, antenna parameters measured by our method share almost no difference from the manual measurement results, which complies with industry standards.

A. SEGMENTATION BASED ON MASK R-CNN

Mask R-CNN is the representative in instance segmentation. After inputting an image into the network, it can output an existing object with high quality mask which is generated for each instance. With the basic framework of Mask R-CNN displayed in Figure 4, what can be seen is that it consists of three core groups: the backbone network, the full convolution network, and the region of interest alignment (ROI Align). The backbone network aims to achieve object detection and classification while the full convolution network, and the region of interest alignment (ROI Align). The backbone network aims to achieve object detection and classification while the full convolutional network is designed to add a mask to the detected object. As for ROI Align, it uses bilinear interpolation to replace traditional quantization operation to reduce error.

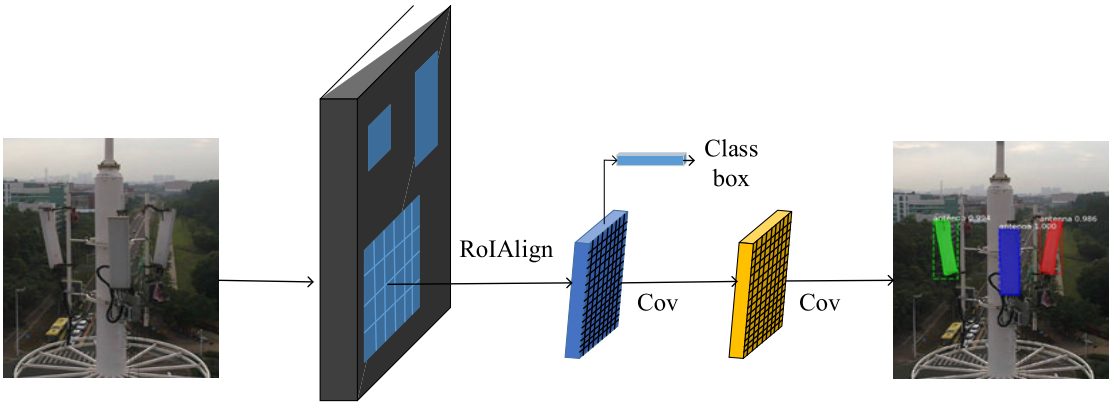


FIGURE 4. Framework of mask R-CNN.

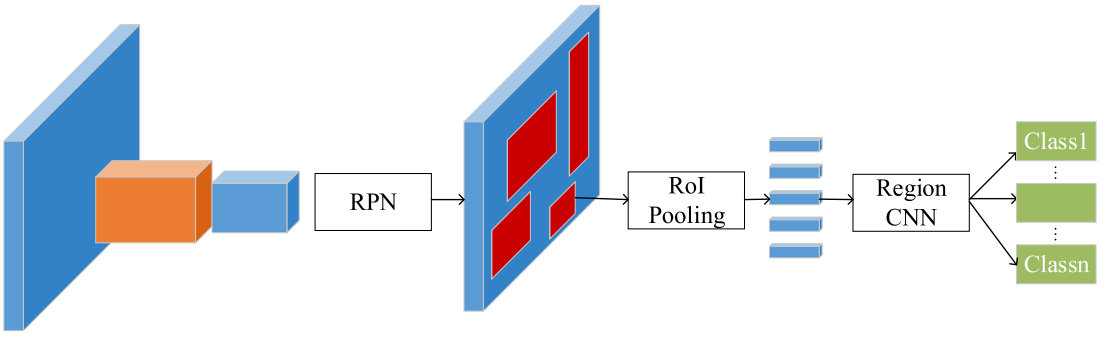


FIGURE 5. Backbone of network of instance segmentation algorithm.

1) BACKBONE NETWORK

Mask R-CNN refers to one of the basic algorithms for object detection, whose backbone network structure is shown in Figure 5. The backbone network devotes to proposing pooling technology of region of interest (ROI) and region proposal network (RPN). RPN enhances detection accuracy while promoting the speed. The role of ROI pooling technique is to fix ROI on feature map to a specific size (7×7) by maximum pooling operation for subsequent classification and bounding box regression operations. However, since the location of pre-selected ROI is usually obtained by model regression, which is generally a floating-point number, ROI pooling technique possesses two processes of data quantification. After the above two quantization operations, there is a certain deviation between the quantified ROI and the original ROI, and this deviation will affect the accuracy of the object detection. RPN is used to distinguish and initially locate multiple ROI generated on feature map. It is robustly implemented in a full convolution manner, using the convolutional feature map returned by the underlying network as input, as well as a convolutional layer with n channels and a 3×3 convolution size and two parallel 1×1 convolution kernels, in which the quantity of channels n shows positive correlation with the number of anchors. At the same time, in the classification layer, the predicted values of the background and objects are outputs for each anchor.

2) FULLY CONVOLUTIONAL NETWORK (FCN)

FCN is a classic network in semantic segmentation, which can accurately segment objects in images. The network framework is manifested in Figure 6 which indicates an end-to-end network. FCN classifies images in pixel-to-pixel manner, thus solving the problem of image segmentation (semantic segmentation) at the semantic level. Varied from the classic CNN after convolutional layer using fully connected layer to obtain fixed-length feature vectors for classification (FC layer + softmax), FCN can accept input images of any size, using the deconvolution layer after last convolution layer to upsample the feature map of the last convolutional layer to restore it to the same size as the input image, so that a prediction can be generated for each pixel while preserving the spatial information in the original input image. Finally, pixel-by-pixel classification is performed on the upsampled feature map and the loss of softmax classification is calculated.

3) ROI ALIGN

Traditional detection framework makes use of ROI Pooling, which pools the corresponding area into a fixed-size feature map in the feature map according to the position coordinates of pre-selected frame for subsequent classification and bounding box regression operations. Since the position of the pre-selected box is usually generated by regression of model, it is generally a floating-point number. Unfortunately,

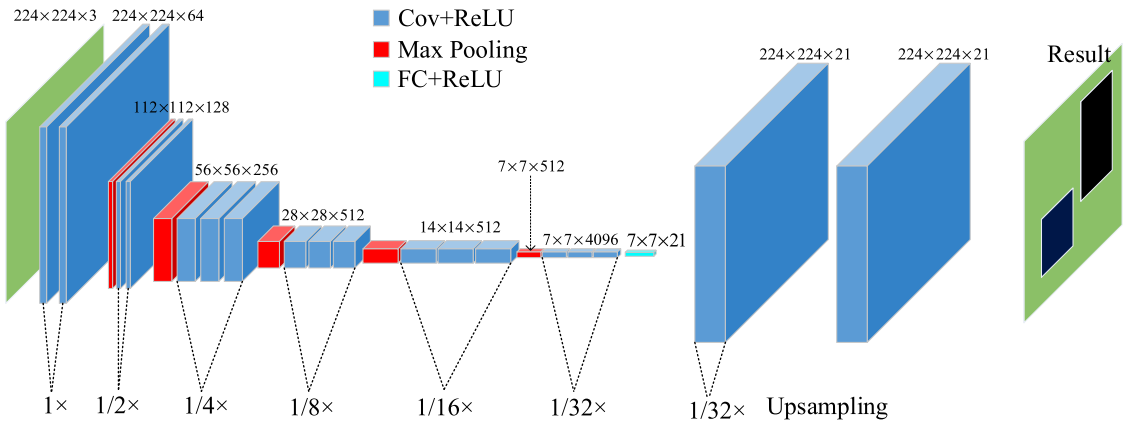


FIGURE 6. Framework of FCN.

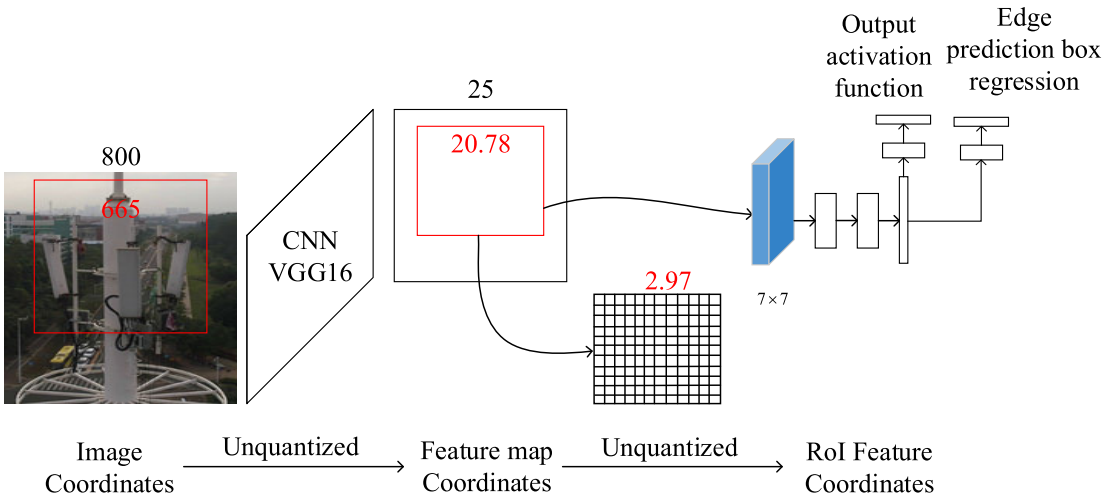


FIGURE 7. Region of interest alignment.

the pooled feature map requires a fixed size. Therefore, the operation of ROI Pooling is equipped with two quantification processes. One is that candidate box boundaries are quantized to integer point coordinate values. The other is that quantized boundary regions are equally divided into k × k bins, and then the boundaries of each cell are quantized. However, after two quantifications, the candidate box has a certain error from the initial position, which affects the accuracy of detection or segmentation. ROI Alignment technique, instead of using the quantization operation in the pool of ROI, employs linear interpolation method whose main purpose is to avoid errors caused by quantization operations. The details of this process are expressed in Figure 7. In order to obtain a stable size (7X7) feature map, ROI alignment technique uses a bilinear interpolation algorithm to process floating point numbers. Bilinear interpolation is a better image scaling algorithm, which fully utilizes four real pixel values around virtual point in the original image to jointly determine a pixel value in target image. This operation can make pixels in the original image and pixels in the feature map completely aligned, without deviation, which not only augments the accuracy of detection, but also facilitates instance segmentation.

B. LINEAR FITTING BASED ON LEAST SQUARES

Least squares method can find the best function match of unknown data by minimizing the sum of the squares of errors. In this paper, least squares method is introduced to optimize the coordinates of the leftmost and rightmost pixel points of the mask generated by Mask R-CNN, fitting curve and coefficient of the best performance, so as to improve the measurement accuracy of antenna parameters, for instance, down-tilt angle. But for the other parameters, the quantization of pixel-level mask can readily meet the need. For example, the area of antenna can be obtained by multiplying the width of antenna mask by the height, and the aspect ratio of antenna can be obtained by dividing the width of antenna mask by the height. Steps of measuring the down-tilt angle of the antenna by least square method are elaborated in the following part.

Let the expression of the line equation be,

$$y = a + bx \tag{1}$$

The best a and b are obtained from the set of pixel points on the right edge of the mask antenna generated by the optimal model. For a couple of pixel point coordinates (x_i, y_i) that satisfy the linear relationship, assumes that the error of the

coordinate x_i of the x-axis of the pixel is negligible. Then under the same x_i, the error d_i of y_i and $a + bx$ is as follows,

$$d_n = y_n - a - bx \tag{2}$$

Due to the fact that optimal model trained by Mask R-CNN cannot reach the extent that all antennas can be detected, and the detected antennas cannot guarantee full padding during the process of adding mask, the set of pixels on the right edge of the mask is generally not in a straight line. In this way, it can be only considered $d_1 + d_2 + \cdots + d_n$ as the minimum, but because $d_1, d_2, \cdots, d_n$ has positive and negative numbers, adding will cancel each other out, Hence, an equivalent method can be taken to eliminate this impact, which is when $d_1^2 + d_2^2 + \cdots + d_n^2$ appears to be the minimum value, $d_1, d_2, \cdots, d_n$ must be the minimum value.

$$D = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} [y_i - a - b_i]^2 \tag{3}$$

D finds the first-order partial derivative of a and b respectively,

$$\frac{\partial D}{\partial b} = -2[\sum_{i=1}^{n} x_i y_i - a \sum_{i=1}^{n} x_1 - b \sum_{i=1}^{n} x_i^2] \tag{4}$$

Find the second-order partial derivative,

$$\frac{\partial^2 D}{\partial b^2} = 2 \sum_{i=1}^{n} x_i^2 \tag{5}$$

Obviously, $\frac{\partial^2 D}{\partial a^2} = 2n \geq 0$; $\frac{\partial^2 D}{\partial b^2} = 2 \sum_{i=1}^{n} x_i^2 \geq 0$, Satisfy the minimum condition, so setting the first-order partial derivative is 0.

$$\sum_{i=1}^{n} x_i y_i - a \sum_{i=1}^{n} x_1 - b \sum_{i=1}^{n} x_i^2 = 0 \tag{6}$$

Average value has been introduced,

$$\overline{xy} = \frac{1}{n} \sum_{i=1}^{n} x_i y_i \tag{7}$$

Then getting the result,

$$a = \bar{y} - b\bar{x} \tag{8}$$

$$b = \frac{\overline{xy} - \overline{xy}}{\overline{x^2} - \bar{x}^2} \tag{9}$$

The values of a and b are substituted into the linear equation $y = a + b$ to obtain a linear regression equation. And calculate the antenna down-tilt angle according to the value of b, Let the antenna down-tilt angle ω_r fitted by the set of pixels on the right edge of the mask. So, there is an equation related ω_r

$$\omega_r = \arctan |b| \tag{10}$$

In order to reduce the error caused by fitting curve of the right edge pixel point, this paper introduces a set of pixels which are the left edge of the full side antenna mask. For the same reason, ω_1 is the down-tilt angle which can be obtained by the left edge pixels of the full side antenna mask.

$$\omega_l = \arctan |b'| \tag{11}$$

Therefore, there is a mobile communication base station antenna down-tilt angle which can be expressed as ω

$$\omega = \frac{\omega_r + \omega_l}{2} \tag{12}$$

The process of the method which combined least squares and Mask R-CNN in antenna parameters measurement is presented in Figure 8. The upper image is the curve fitted by all the edge pixels of the antenna mask while the bottom one is the curve fitted to the pixel edge on the right edge of the antenna mask.

C. AUTOMATIC FITTING AND MEASURING

Simultaneously, a fully automatic antenna parameters measurement system using UAV has been proposed which is described in Algorithm 1 and its performance exceeds the state-of-the-art methods. As what have mentioned, semi-automatic antenna parameters measurement is based on Mask R-CNN, least squares and UAVs. The difference between fully automatic method and semi-automatic method lies in that video framing, determining quantity of antennas, finding side antenna image and outputting their parameters. In this process, pixel coordinates and threshold are utilized for measuring antenna quantity and separate all antenna data thereby achieving fully automatic antenna parameters measurement. The automatic antenna parameters measuring algorithm details are shown as below. According to the requirements of UAVs flying antenna mentioned in Session 3.1, Firstly, a complete video of base station antenna taken by UAV is used as the input of the proposed method; Secondly, the video is framed according to the requirements of 15 frames per second and each frame of antenna image is saved together. Then, these original images are input to the optimal model which is trained Mask R-CNN on UAV-Antenna, and its output images with antenna mask are shown in Figure 9.

There is an analyze of antenna images with masks, assuming that the pixel coordinates of the same position of the similar position antenna do not exceed 50 pixel values in the two frames before and after, the two antennas can be considered as one same antenna. Conversely, two different antennas are recorded. Due to the dearth of training samples, the method proposed in this paper may mistake non-antenna objects as antennas, leading to false detection. In order to reduce the impact of the false detection, this paper sets the threshold of the effective frame sequence length at 400 which is the empirical value obtained after many experiments. When the antenna effective frame sequence length is greater than 400, the detected antenna is considered to be a real antenna; otherwise, a false detection caused by other objects in the background like buildings. The whole process is expressly demonstrated in Figure 10.

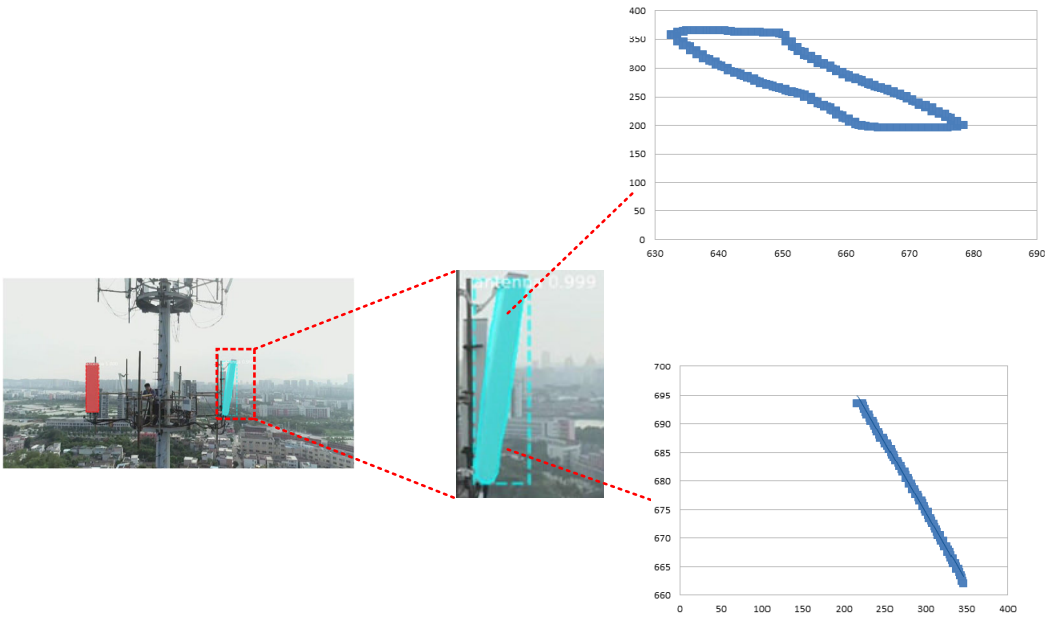


FIGURE 8. Fitting the antenna down-tilt angle.

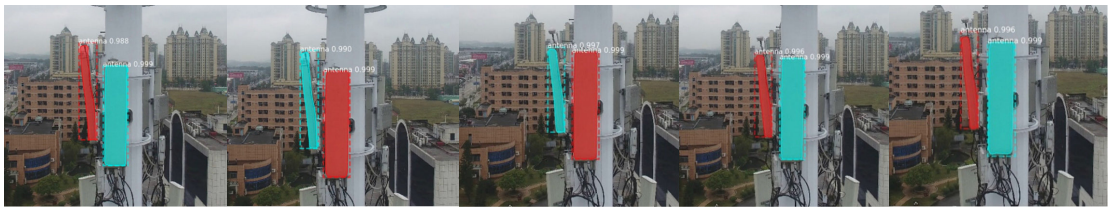


FIGURE 9. Recognition result of 5 consecutive frames from UAV antenna video.

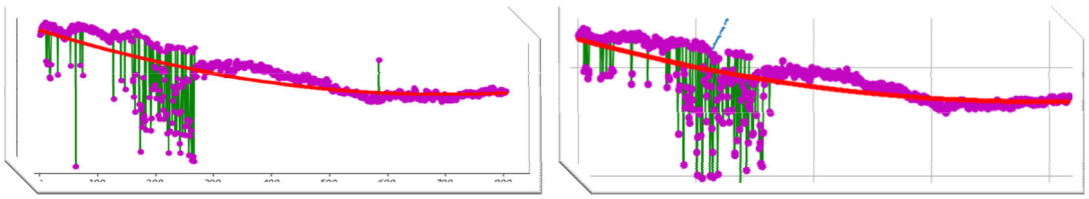


FIGURE 10. Verify and separate the number of antennas by different parameters.

Motivated by the above advancement, data fitting is performed on antenna down-tilt angle, the area of antenna mask, and the aspect ratio of antenna mask to verify and separate the number of antennas in all output images which is shown in Figure 11. At the same time, the parameters which include down-tilt, area, aspect ratio, direction angle (provided by the UAV) and GPS are displayed on our APP terminal interface of all antennas of the mobile communication base station. Figure 12 can be referred to the whole APP terminal interface.

V. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL PARAMETER SETTING

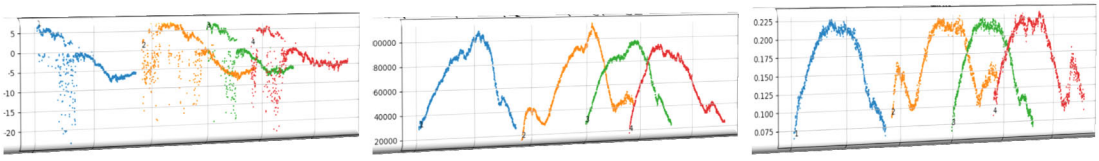
The computer used in experiments is configured with Xeon E3 CPU, NVIDIA GeForce GTX 1080, and 64G memory. The experiments were carried out with ubuntu16.04 operating system, CUDA Toolkit 8.0, and Tensorflow framework. Specific distribution of UAV-Antenna is shown in Table 1. Angles (1°, 4°, 5°, 6°, 7°, 8°, 9°, 12°, 15°) of antennas in the testing set were adjusted by professional workers.

TABLE 1. Specific distribution of UAV-antenna.

Training Set (Unlabeled)	Testing Set (Down-tilt angle label distribution)								
19496	1	4	5	6	7	8	9	12	15
	13	25	22	30	25	23	22	39	20

B. EXPERIMENTAL PARAMETER SETTING

In order to verify the accuracy of proposed antenna parameters measurement method, experiments were conducted on



(a) down-tilt angle (b) area (c) aspect ratio

FIGURE 11. Process of determining the number of antennas.

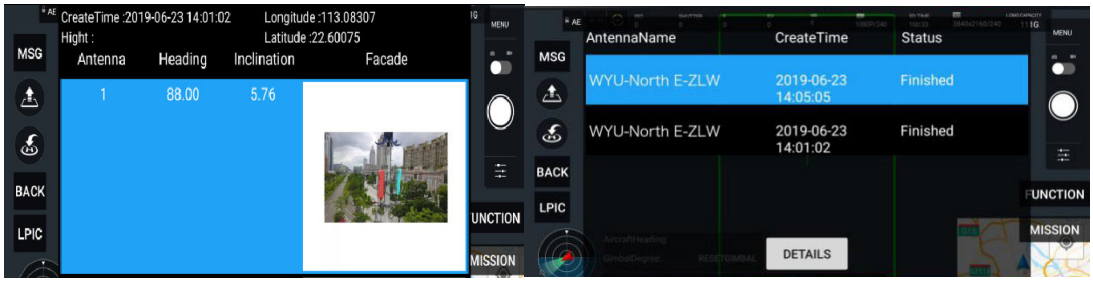


FIGURE 12. APP terminal interface.

Algorithm 1 Antenna Parameters Automatic Measuring

Input:

Video data A of base station through HD cameras provided on UAVs

output:

The quantity n of antenna and the parameters p of each antenna

Step1. Video data A is framed into testing set a at f frames per second, images in a are sorted in chronological order;

Step2. Testing set a is tested by the optimal model trained by Mask R-CNN and generate visualize images;

Step3. Output pixel coordinates of the upper right corner of all masked antennas, pixel interval x and threshold t are utilized to determine the quantity n of antenna;

Step4. Find the image of the side of the antenna based on the image of the largest area of antenna, then output antenna down-tilt according to least squares method via equation (1)-(12);

Step5. Output the quantity n of antenna and the parameters vector p, including down-tilt angle area, aspect ratio, direction angle of each antenna.

training set with 19,496 unlabeled images and tested on testing set with 219 labeled images. After experimental comparison, when learning rate was set at 0.001 and threshold was set at 0.85, the objects detection performed best. Under the premise that learning rate and threshold are unchanged, we conducted several experiments and defined detection accuracy, fitting accuracy and time via optimizing parameters such as epochs, training layers shown in Table 2. Detection accuracy indicates detection rate of all antennas (except the back antenna) in testing set, while fitting accuracy represents

TABLE 2. Performance comparison of testing on various parameters.

Epochs	Layers		Detection Accuracy	Fitting Accuracy	Times(s)
	All	Heads			
30	Y	N	99.3%	46.4%	6.05
30	N	Y	77.2%	36.2%	6.42
40	Y	N	99.4%	48.1%	6.00
40	N	Y	86.7%	29.7%	5.95
50	Y	N	**99.4%**	**58.2%**	**5.95**
50	N	Y	90.2%	47.4%	6.31

ratio of antennas that match the industry standard to the detected antenna, and time is expressed as the time of detecting and fitting parameters (down-tilt angle, area, aspect ratio). By comparing time, detection accuracy and fitting accuracy, optimal model is spotted whose detection accuracy is 99.4%, fitting accuracy is 58.2% and time is 5.95s.

By experiments, when epochs are set at 50 and layers are controlled at all, we can get the best experimental results in terms of detection accuracy, fitting accuracy and time. Hence, this paper selects this model with the best performance to measure antenna images at various angles in the test set. The results can be observed in Table 3. Among them, the deviation angle is the absolute value of the measurement result and the actual angle. It can be known from the deviation angle that the accuracy of mobile communication base station antenna parameters measurement result of the optimal model ranks the top.

The widely used YOLOv3, Faster R-CNN and Mask R-CNN are compared on UAV-Antenna database. Table 4 shows their detection accuracy respectively. Data conclude that Mask R-CNN achieves the best results: 99.44%. Mask R-CNN has less localization error, indicating that Mask R-CNN can localize objects better in terms of antennas detection task.

TABLE 3. Antenna angle measurement result of test set.

Actual Angle	1	4	5	6	7
Measure Angle	2.34	3.50	5.16	5.46	6.01
Deviation	1.34	0.50	0.16	0.54	0.99
Actual Angle	8	9	12	15	7
Measure Angle	7.39	8.77	10.6	13.91	6.01
Deviation	0.61	0.23	1.40	1.09	0.99

TABLE 4. Performance comparison of object detection accuracy.

Method	Detection Accuracy
YOLOv3[15]	92.72%
Faster R-CNN[16]	89.85%
Mask R-CNN[32]	**99.44%**

Thus, using Mask R-CNN to detect the antenna of mobile communication base station proves to be the best choice. Figure 2 is the test results of testing antenna images using the proposed method. The upper images are the original images, while the bottom images are the visualized images obtained from optimal model. The classification results of antennas and accuracy of bounding box are displayed in the visualization images.

VI. CONCLUSION

An intelligent and fully automatic antenna parameters measurement method is successfully proposed using UAVs for mobile communication base station. Our method emphasizes the united collaboration of Mask R-CNN, least squares, frame sequence analysis and UAV to fully automatically measure antenna parameters, skillfully realizing multi-field cooperation of software and hardware. It has been experimentally validated that the appropriate training strategies employed can achieve outstanding performance in fully automatic antenna parameters measurement. Compared with traditional method and the most advanced methods including YOLOv3 and faster R-CNN which have been trained and tested on UAV-Antenna database constructed by image capturing with the help of UAVs flying around various base stations in antenna parameter measurement, our method significantly excels in its features as low cost, low dependence on hardware methods and easiness in application. In addition, the innovated proposed method has filled the blank in the field of previous antenna parameters measurement methods using algorithms coupled with UAV. Simultaneously, it produces markedly superior measurement efficiency and satisfactory safety control.

It is believed that our method, apart from its standalone utility, provides a useful solution for mobile communication base station antenna parameters measurement. It stands an extraordinarily promising future for it is a constructive system which used CNN to detect antennas and measured its parameters in more complex scenarios or any other low-altitude object detection and parameters measurement in video simultaneously.

REFERENCES

[1] A. Geise, O. Neitz, J. Migl, H.-J. Steiner, T. Fritzel, C. Hunscher, and T. F. Eibert, "A crane-based portable antenna measurement system—System description and validation," *IEEE Trans. Antennas Propag.*, vol. 67, no. 5, pp. 3346–3357, May 2019.

[2] M. Garcia-Fernandez, Y. Alvarez-Lopez, F. L. Heras, A. Arboleya-Arboleya, B. Gonzalez-Valdes, and Y. Rodriguez-Vaqueiro, "Unmanned aerial system for antenna measurement (UASAM)," in *Proc. Eur. Conf. Antennas Propag.*, 2018, pp. 794–796.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[4] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, 2014.

[5] S. Levine, P. Pastor, and A. Krizhevsky, "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *Int. J. Robot. Res.*, vol. 37, nos. 4–5, pp. 421–436, 2018.

[6] S. Dörner, S. Cammerer, J. Hoydis, and S. ten Brink, "Deep learning based communication over the air," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 132–143, Feb. 2018.

[7] M. G. Fernandez, Y. A. Lopez, and F. L.-H. Andres, "On the use of unmanned aerial vehicles for antenna and coverage diagnostics in mobile networks," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 72–78, Jul. 2018.

[8] H. Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, and M. Guizani, "Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges," *IEEE Access*, vol. 7, pp. 48572–48634, 2019.

[9] M. Shahzad, M. Maurer, F. Fraundorfer, Y. Wang, and X. X. Zhu, "Buildings detection in VHR SAR images using fully convolution neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1100–1116, Feb. 2019.

[10] G. Cheng, J. Han, P. Zhou, and D. Xu, "Learning rotation-invariant and Fisher discriminative convolutional neural networks for object detection," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 265–278, Jan. 2019.

[11] A. Raj, K. Gandhi, B. T. Nalla, and N. K. Verma, "Object detection and recognition using small labeled datasets," in *Computational Intelligence: Theories, Applications and Future Directions—Volume II* (Advances in Intelligent Systems and Computing), vol. 799, N. Verma and A. Ghosh, Eds. Singapore: Springer, 2019.

[12] D. Zhang, J. Han, L. Zhao, and D. Meng, "Leveraging prior-knowledge for weakly supervised object detection under a collaborative self-paced curriculum learning framework," *Int. J. Comput. Vis.*, vol. 127, no. 4, pp. 363–380, 2019.

[13] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published.

[14] Y. Zhou, A. Mao, S. Huo, J. Lei, and S. Kung, "Salient object detection via fuzzy theory and object-level enhancement," *IEEE Trans. Multimedia*, vol. 21, no. 1, pp. 74–85, Jan. 2019.

[15] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1440–1448.

[16] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Computer Vision—ECCV* (Lecture Notes in Computer Science), vol. 9905, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham, Switzerland: Springer, 2016.

[17] 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.

[18] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.

[19] W. Xiang, H. Mao, and V. Athitsos, "ThunderNet: A turbo unified network for real-time semantic segmentation," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Jan. 2019, pp. 1789–1796.

[20] S. Gharghabi, C.-C. M. Yeh, Y. Ding, W. Ding, P. Hibbing, S. LaMunion, A. Kaplan, S. E. Crouter, and E. Keogh, "Domain agnostic online semantic segmentation for multi-dimensional time series," *Data Mining Knowl. Discovery*, vol. 33, pp. 1–2, Jan. 2019.

[21] Y. Zhang, P. David, H. Foroosh, and B. Gong, "A curriculum domain adaptation approach to the semantic segmentation of urban scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, to be published.
[22] X. Zhang, Z. Chen, Q. M. J. Wu, L. Cai, D. Lu, and X. Li, "Fast semantic segmentation for scene perception," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 1183–1192, Feb. 2019.
[23] G. Ciocca, D. Mazzini, and R. Schettini, "Evaluating CNN-based semantic food segmentation across illuminants," in *Computational Color Imaging—CCIW* (Lecture Notes in Computer Science), vol. 11418, S. Tominaga, R. Schettini, A. Trémeau, and T. Horiuchi, Eds. Cham, Switzerland: Springer, 2019.
[24] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 801–818.
[25] P. Wang, P. Chen, Y. Yuan, D. Liu, Z. Huang, X. Hou, and G. Cottrell, "Understanding convolution for semantic segmentation," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2018, pp. 1451–1460.
[26] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018.
[27] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
[28] D. Ikami, T. Yamasaki, and K. Aizawa, "Fast and robust estimation for unit-norm constrained linear fitting problems," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8147–8155.
[29] N. B. Kar, K. S. Babu, A. K. Sangaiah, and S. Bakshi, "Face expression recognition system based on ripplet transform type II and least square SVM," *Multimedia Tools Appl.*, vol. 78, no. 4, pp. 4789–4812, 2019.
[30] K. Kumar, "Discrete wavelet transform (DWT) assisted partial least square (PLS) analysis of excitation-emission matrix fluorescence (EEMF) spectroscopic data sets: Improving the quantification accuracy of EEMF technique," *J. Fluorescence*, vol. 29, no. 1, pp. 185–193, 2019.
[31] A. Ferragina, F. Benozzo, and P. Berzaghi, "Bayesian and partial least square global forage calibrations models developed by an iterative procedure using R," in *Proc. 18th Int. Conf. Near Infr. Spectrosc.*, 2019, pp. 51–55.
[32] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2961–2969.
[33] C. Niu, Q. Yang, S. Ren, H. Hu, D. Han, Z.-J. Hu, and J. Liang, "Instance segmentation of auroral images for automatic computation of arc width," *IEEE Geosci. Remote Sens. Lett.*, to be published.
[34] K. Halupka, R. Garnavi, and S. Moore, "Deep semantic instance segmentation of tree-like structures using synthetic data," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Jan. 2019, pp. 1713–1722.
[35] T.-N. Le and A. Sugimoto, "Semantic instance meets salient object: Study on video semantic salient instance segmentation," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Jan. 2019, pp. 1779–1788.

YIKUI ZHAI received the bachelor's degree in optical electronics information and communication engineering and the master's degree in signal and information processing from Shantou University, Guangdong, China, in 2004 and 2007, respectively, and the Ph.D. degree in signal and information processing from Beihang University, in June 2013. He has been a Visiting Scholar with the Department of Computer Science, Universitá degli Studi di Milano, since 2016. He is currently an Associate Professor with Wuyi University, Guangdong, China, where he has been with the Department of Intelligence Manufacturing, since 2007. His research interests include image processing, deep learning, and pattern recognition.

QIRUI KE received the B.S. degree from the Hubei University of Arts and Science, in 2018. He is currently pursuing the master's degree with the Department of Intelligence Manufacturing, Wuyi University. His research interests include biometric extraction and pattern recognition.

YING XU received the B.S. degree in automation and the M.S. degree in control engineering from the Wuhan University of Science and Technology, in 2004 and 2008, respectively, and the Ph.D. degree from the South China University of Technology, in 2013. She joined Wuyi University, in 2008. Her research interests include intelligent signal processing and pattern recognition.

WENBO DENG received the B.S. degree from the Hubei University of Arts and Science, in 2016. He is currently pursuing the master's degree from the Department of Intelligence Manufacturing, Wuyi University. His research interests include biometric extraction and pattern recognition.

JUNYING GAN received the B.S., M.S., and Ph.D. degrees in electrical information engineering from Beihang University, in 1987, 1992, and 2003, respectively. She is currently a Full Professor with Wuyi University, Guangdong, China. She joined Wuyi University, in 1992. She is the Executive Director of the Guangdong Image Graphics Association. She has published over 50 journal papers. Her research interests include biometric extraction and pattern recognition. She has received several provincial technology awards.

JUNYING ZENG was born in Ganzhou, China, in 1977. He received the Ph.D. degree in physical electronics from the Beijing University of Posts and Telecommunications, Beijing, China, in 2008. His research interests include intelligent signal processing and pattern recognition.

WENLVE ZHOU received the B.S. degree from the Beijing Institute of Technology, Zhuhai, in 2019. He is currently pursuing the master's degree with the Department of Intelligence Manufacturing, Wuyi University. His research interests include biometric extraction and pattern recognition.

FABIO SCOTTI received the Ph.D. degree in computer engineering from the Politecnico di Milano, Milan, Italy, in 2003. He has been an Associate Professor in computer science with the Universitá degli Studi di Milano, Crema, Italy, since 2015. His original results have been published in more than 100 papers in international journals, proceedings of international conferences, books, book chapters, and patents. His current research interests include biometric systems, machine learning and computational intelligence, signal and image processing, theory and applications of neural networks, three-dimensional reconstruction, industrial applications, intelligent measurement systems, and high-level system design. He is an Associate Editor of the IEEE TRANSACTIONS ON HUMAN–MACHINE SYSTEMS AND SOFT COMPUTING (Springer). He has been an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY and a Guest Coeditor of the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT.

RUGGERO DONIDA LABATI received the Ph.D. degree in computer science from the Universitá degli Studi di Milano, Crema, Italy, in 2013, where he has been an Assistant Professor in computer science, since 2015. He has been a Visiting Researcher with Michigan State University, East Lansing, MI, USA. His original results have been published in more than 50 papers in international journals, proceedings of international conferences, books, and book chapters. His current research interests include intelligent systems, signal and image processing, machine learning, pattern analysis and recognition, theory and industrial applications of neural networks, biometrics, and industrial applications. He is an Associate Editor of *the Journal of Ambient Intelligence and Humanized Computing* (Springer).

VINCENZO PIURI received the M.S. and Ph.D. degrees in computer engineering from the Politecnico di Milano, Italy. He has been a Full Professor with the Universitá degli Studi di Milano, Italy, since 2000, where he was also the Department Chair, from 2007 to 2012. He was an Associate Professor with the Politecnico di Milano, Italy, from 1992 to 2000, a Visiting Professor with the University of Texas at Austin, USA, from 1996 to 1999, and a Visiting Researcher with George Mason University, USA, from 2012 to 2016. He founded a start-up company, Sensure srl, in the area of intelligent systems for industrial applications (leading it from 2007 to 2010). He was active in industrial research projects with several companies. His main research and industrial application interests include intelligent systems, computational intelligence, pattern analysis and recognition, machine learning, signal and image processing, biometrics, intelligent measurement systems, industrial applications, distributed processing systems, the Internet-of-Things, cloud computing, fault tolerance, application-specific digital processing architectures, and arithmetic architectures. He is an ACM Fellow.