A prohibited object detection and recognition system for X-ray images

Yang Zheng, Jie Liu, Shuwu Zhang
Institute of Automation Chinese Academy of Sciences
Beijing, China
yang.zheng@ia.ac.cn

2021 International Conference on Culture-oriented Science & Technology (ICCST) | 978-1-6654-4254-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCST53801.2021.00115

***Abstract*—The object detection and recognition system for X-ray images containing prohibited object divided into two parts: The construction of synthetic data containing prohibited objects, as well as the detection and recognition of prohibited objects. Among them, the construction of prohibited object data includes object classification, attribute analysis, object image acquisition, fusion of object image and X-ray image. After the above steps, the X-ray images with prohibited objects such as knives and guns are similar to the real images are obtained, which effectively solves the problem of insufficient X-ray image. In the process of detection and recognition of prohibited objects, by changing the color and shape of training data, the data source is expanded, which makes the system have higher robustness and improves the detection efficiency of prohibited objects. The detection and recognition system can effectively complete the real-time detection and recognition of knives and guns in luggage, save a lot of manpower and material resources, and ensure the safety of road, railway and air transportation.**

Keywords- X-ray images; prohibited object; detection and recongition

I. INTRODUCTION

Image monitoring of X-ray security machines requires a large number of staff to determine objects in real time from the X-ray image configuration of X-ray security machines. In order to improve the speed of security screening, reduce passenger waiting time, as well as save a lot of manpower and material resources, it is necessary to upgrade the monitoring system to improve the intelligence of the system. Today's popular object detection method is a deep learning method in artificial intelligence, which requires a large amount of image data. In order for the model to detect different sizes, different angles, different brands, different categories of knives and guns, a large number of X-ray images containing prohibited objects are required.

In order to meet the requirements of artificial intelligence model large quantities of data, a large number of knives and guns to be placed in different packages, X-ray machine processing. In the real world, because knives and guns are controlled objects, and a wide variety of brands, so through the security X-ray machine to obtain image data methods are difficult to achieve. Therefore, it is necessary to obtain more comprehensive image data through other methods to meet the needs of system upgrades.

In the X-ray image, the appeared various colors of objects are a reflection of the density, quality and quantity of objects under X-ray. Security X-ray machine can store a large number of X-ray images without knives and guns, and it can download images of different types of knives and guns on the Internet at the same time. Therefore, according to the principle that different objects in the X-ray machine image present different colors, the X-ray machine image containing different sizes, shapes and categories of knives and guns can be generated by fusing the knife and gun images and X-ray images in the network.

X-ray security image processing method is different from the conventional method, its own imaging principle, acquisition methods, etc. can lead to data bias. Because of their different penetration capability, the X-ray image data taken by different X-ray machine manufacturers will have some deviation in the representation and characteristic distribution. At the same time, prohibited objects in X-ray images show different sizes and shapes, increasing the difficulty of detection and recognition. In order to make the system have a high robustness, in the process of training the deep learning model, the data expansion is completed by changing the shape, color and other properties of prohibited objects.

The common object detection algorithm is mainly divided into SSD[1] and Yolo series[2,3,4,5] as the representative of the one step object detection algorithm and Faster-Rcnn[6] as the representative of the two steps object detection algorithm, because the detection and recognition system of prohibited objects has a higher requirements for the real-time detection, so the one step object detection algorithm YOLO as the basis of the object detection algorithm. In order to be able to have a high recall rate for prohibited objects, and to avoid missing any suspicious objects, a lower threshold is used.

The remainder of the paper is organized as follows: The detection and recognition system is described in Section 2. Experiments are presented in Section 3.

II. APPROACH

The detection and recognition system of X-ray images consists of two parts: the construction of synthetic data

978-1-6654-4254-1/21/$31.00 ©2021 IEEE
DOI 10.1109/ICCST53801.2021.00115

containing prohibited objects and the detection and recognition of prohibited objects. The prohibited objects mean knives and guns in the paper. As shown in Figure 2-1, the details will be described.

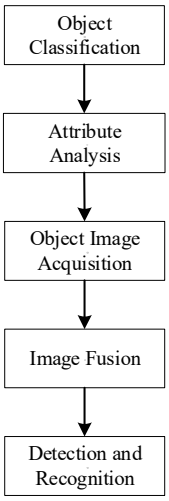

Figure 2-1 Chart of detection and recognition system

2.1 Object Classification

Different kinds of objects show different shapes in X-ray images. At the same time, different individuals in the same category also have differences. In order to generate the image more in line with the requirements, the categories of prohibited objects are distinguished in detail, which helps to set different parameters according to different categories in the fusion process.

According to the different sizes and shapes of knives and guns, the categories of knives and guns are divided into pistols, submachine guns, long guns, daggers, kitchen knives, long knives and other categories.

2.2 Audio Fingerprint Extraction

There are some X-ray images in the network containing different categories. Among them, the SIXray image data published by the team of the University of Chinese Academy of Sciences[7] and the data published by iFLYTEK are consistent with the pseudo color images scanned by the security X-ray machine, and the images are more in line with the actual needs.

The knife and gun in each X-ray image are analyzed respectively, and their attributes such as the length and width in the image, the proportion in the image, the location in the image and the color information in RGB channels are calculated. According to the analysis of X-ray image, the size, proportion, pixel value and other specific information of pistol, submachine gun, long gun, dagger, kitchen knife, long knife and other categories of individuals in the image are calculated. For example, the length and width of pistol, dagger and kitchen knife compared with the length and width of the trunk will not exceed 0.3, and the color is mainly blue. Some examples are shown in Figure 2-2.

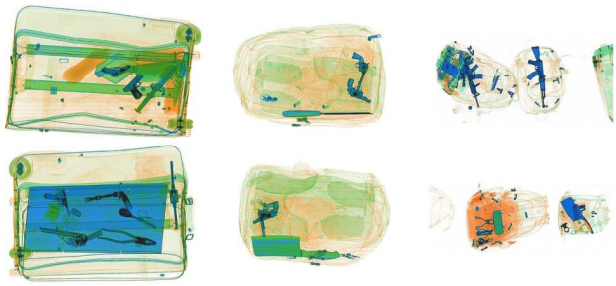

Figure 2-2 X-ray images including pistol、kitchen knife、submachine gun

2.3 Object Image Acquisition

Using the method of web crawler, the images of knives and guns with different sizes, types and shapes from different professional material websites can be downloaded: (1) Confirm the address of the professional material website, and use the python requests network request package to send the request to the object address; (2) The request returns the source code of the web page, analyzes the structure of the web page with XPath or beautiful soup, and parses the links of each image; (3) Send a request for each image link, write the returned result to the local, and store it as an image file; (4)Delete the image that does not meet the requirements. Some knives and guns are shown in Figure 2-3.

Figure 2-3 Pistols、submachine guns、long guns、 daggers、 kitchen knives、long knives，downloaded from the network.

By analyzing and comparing the images of knives and guns obtained from professional material websites, the training data of convolution neural network model is determined. According to the requirements of generating knives and guns and reference to the network structure of the Generative Adversarial Networks[8], the generation model is designed. The training data of the model is mainly composed of the images downloaded from the network, with a total of 22006 images. By training the data, the required anti-network model is obtained, and the knives and guns images are expanded.

2.4 Fusion of Object Image and X-ray Image

Using the methods of image processing, the knives and guns are incorporated into the X-ray image according to certain rules and principles, the X-ray image containing the prohibited objects is obtained, and the location information of the prohibited objects in the image is stored, detailed steps are as follows:

Firstly, the pixel value range of the knife and gun in the RGB channels of the X-ray image is determined. Since the objects such as the knife and the gun will show different shades of blue in the pseudo color image of the X-ray machine, the pixel value of the B channel is fixed in a certain range.

Secondly, in order to reflect the fusion effect of X-ray image and object image, the pixel value of X-ray image and object image in the R-channel and G-channel of fusion area accounts for a certain proportion, that is, set the corresponding threshold for fusion.

$$\begin{cases} Blue1_{value} \le Blue_{value} \le Blue2_{value} \\ Green_{value} = thred1 \times Xray_{greenValue} + thred2 \times Object_{greenValue} \\ \mathrm{Re}d_{value} = thred3 \times Xray_{redValue} + thred4 \times Object_{redValue} \end{cases} \quad (2\text{-}1)$$

The formula (2-1) is the calculation method of region fusion. $Blue_{value}$ 、 $Green_{value}$ 、 $\mathrm{Re}d_{value}$ are the pixel value after fusion. $Blue1_{value}$ 和 $Blue2_{value}$ are in the range of channel B for the region fusion. Generally, the value ranges from 160 to 210. $Xray_{greenValue}$ 、 $Xray_{redValue}$ 、 $Object_{greenValue}$ and $Object_{redValue}$ respectively represent the pixel values of X-ray images and objects in G channel and R channel, $thred1$ 、 $thred2$ 、 $thred3$ and $thred4$ respectively account for the proportion of pixel values of each image in the fusion area. The parameters in the formula are calculated by the real X-ray images.

Finally, the area of the object image is incorporated into the X image and the specific location information is stored. The process of generating X-ray image containing knives and guns shown in Figure 2-4.

The X-ray image containing knives and guns is generated, and the specific position of the corresponding object in the image is stored.

2.5 Detection and Recognition

In order to make the detection and recognition model have generalization performance, the X-ray images containing prohibited objects are exposed and cropped, and the color values are adjusted in HSV space to make the X-ray images present different brightness, size and color. The operation of X-ray image in HSV space can converted to different color images. The process is shown in Figure 2-5 and the details are as follows:

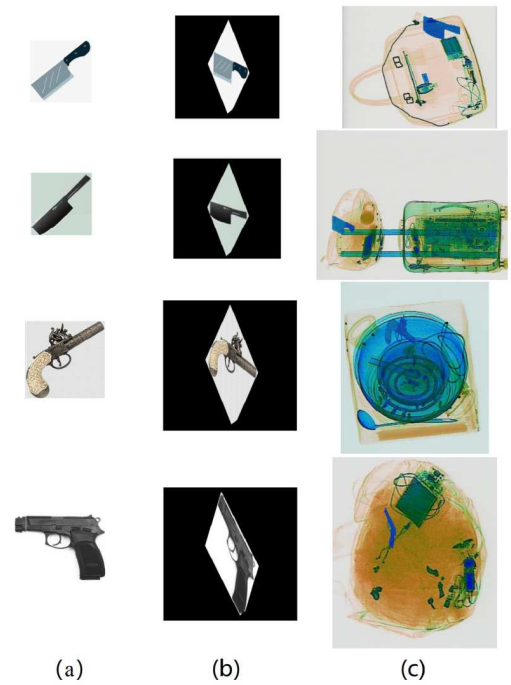

Figure 2-4 The process of generating X-ray images containing knives and guns. (a) Images of knives and guns obtained from the network; (b) The area to be inserted into the X-ray image is the foreground area; (c) According to a certain principle, the fusion result of the image of the prohibited objects and the X-ray image is obtained.

Step 1: The X-ray image is transformed from RGB space to HSV space by using OPENCV functions;

Step 2: Adjust the hue H, saturation S and brightness V of the image in HSV space respectively, so that it has obvious changes with the HSV value of the original image;

Step 3: The image is converted from HSV space to RGB space by using OPENCV functions, and the image is stored.

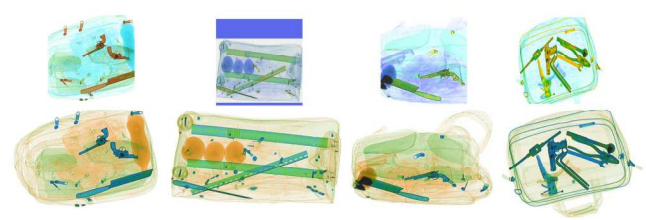

Figure 2-5 Some examples of X-ray images containing knives and guns and the result after clipping, scaling and HSV transformation.

The structure of the YOLO model with the function of detecting and recognizing prohibited objects is defined. Convolution neural network model is used to detect and recognize the X-ray images, and a high

detection accuracy of prohibited objects is obtained. According to the category of X-ray image detection and recognition and the size of training data, the parameters in the network model are set, and the data are trained by graphics card to get the detection and recognition network model.

The trained model is used to detect and recognize the X-ray image, and the coordinate and category data output by the model are stored. So far, we have completed the location and category judgment of the prohibited objects such as knives and guns in the X-ray image.

III. EXPERIMENTS

In order to prove the validity of the synthetic data, this paper respectively carries out the detection and recognition on the real data and the synthetic data. The real data comes from the University of Chinese Academy of Sciences, iFLYTEK and other websites in the network, with a total of 4631 images, including 4031 training images and 600 testing images. The synthetic data contains 6018 images for training, and 800 images for testing.

The hardware of this paper is 2080TI graphics card. The initial parameters of the model are the parameters obtained from 80 categories of COCO data[9]. The bath size is set to 2, and the number of iteration is set to 10000. The size of prohibited objects in X-ray image is different from that of 80 categories in COCO data. Therefore, the size of object data in X-ray image is clustered to satisfy the requirement.

Table 3-1 Performance of different dataset

Training data	Testing data	Recall(%)	Precison(%)
Real data	Real data	70	75
	Synthetic data	76	79
Real data Synthetic data	Real data	75	79
	Synthetic data	92	91

By analyzing of the table 3-1, it can be seen that due to the variety and shape of the prohibited object in the real data, and the complex background, the detection and recognition result of the prohibited object in the real data is lower. When the real data and the synthetic data are used as the training data, the recall and precision are higher than that of the single data set as the training set. It can be seen that the synthetic data has a certain auxiliary effect, and some detection results are shown in the figure 3-1.

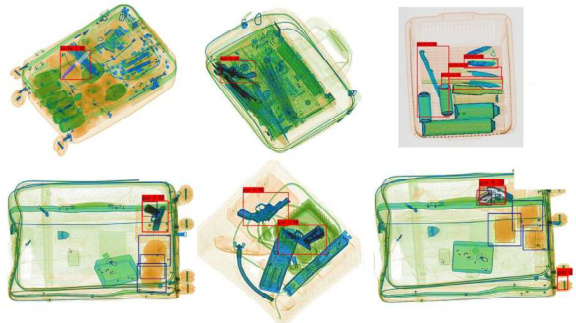

Figure3-1 The results of detection and recognition.

IV. CONCLUSION

The proposed X-ray machine detection and recognition system for prohibited object such as knives and guns can better assist the security personnel to complete the detection and recognition. However, due to the complexity of the real environment, it is difficult to achieve the same performance as the real data. Therefore, the way how to further construct more synthetic data is the focus of future research. At the same time, it is very important to further optimize the detection and recognition model.

ACKNOWLEDGEMENTS

This work was supported by the National Key R&D Program of China (2018YFC0809200). It was also the research achievement of the Key Laboratory of Digital Rights Services.

REFERENCES

[1] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision. Springer, Cham, 2016.

[2] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[J]. IEEE, 2016.

[3] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[J]. IEEE Conference on Computer Vision & Pattern Recognition, 2017:6517-6525.

[4] YOLOv3: An Incremental Improvement[J]. arXiv e-prints, 2018.

[5] Bochkovskiy A, Wang C Y, Liao H. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. 2020.

[6] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149

[7] Miao C, Xie L, Wan F, et al. SIXray: A Large-Scale Security Inspection X-Ray Benchmark for Prohibited Item Discovery in Overlapping Images[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

[8] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks[J]. Advances in Neural Information Processing Systems, 2014, 3:2672-2680

[9] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Contest[J]. Springer International Publishing, 2014