IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON AI-DRIVEN BIG DATA PROCESSING: THEORY, METHODOLOGY, AND APPLICATIONS

Received January 31, 2019, accepted February 19, 2019, date of publication February 28, 2019, date of current version March 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2902121

Data Augmentation for X-Ray Prohibited Item Images Using Generative Adversarial Networks

JINFENG YANG[1,2], ZIHAO ZHAO[1], HAIGANG ZHANG[1], AND YIHUA SHI[1]
[1]Tianjin Key Laboratory for Advanced Signal Processing, Civil Aviation University of China, Tianjin 300300, China
[2]Shenzhen Polytechnic, Shenzhen 518055, China

Corresponding author: Jinfeng Yang (jfyang@cauc.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61806208 and Grant 61379102.

ABSTRACT Recognizing prohibited items automatically is of great significance for intelligent X-ray baggage security screening. Convolutional neural networks (CNNs), with the support of big training data, have been verified as the powerful models capable of reliably detecting the expected objects in images. Therefore, building a specific CNN model working reliably on prohibited item detection also requires large amounts of labeled item image data. Unfortunately, the current X-ray baggage image database is not big enough in count and diversity for CNN model training. In this paper, we propose a novel method for X-ray prohibited item data augmentation using generative adversarial networks (GANs). The prohibited items are first extracted from X-ray baggage images using a K-nearest neighbor matting scheme. Then, the poses of the obtained item images are estimated using a space rectangular coordinate system and categorized into four or eight classes for constructing a training database. For generating the realistic samples reliably, different GAN models are evaluated using Frechet Inception Distance scores, and some important tips of handling GAN training in X-ray prohibited item image generation are reported. Finally, to verify whether the generated images belong to its corresponding class or not, a cross-validation scheme based on a CNN model is implemented. The experimental results show that most of the generated images can be classified correctly by the CNN model. This implies that the generated prohibited item images can be used as the extended samples to augment an X-ray image database.

INDEX TERMS Data augmentation, generative adversarial networks, image generation, prohibited item, X-ray image.

I. INTRODUCTION

X-ray security baggage screening systems are widely installed in almost every station/airport to ensure public transport security [1]. But the reliability of manual detection have been undesirable in real situation. For a baggage screener, detecting the prohibited items is usually so laborious and boring that missing some threatening items is unavoidable in practice [2]. Especially, in rush hours, passengers usually take a lot of time waiting for security checking in line. A reliable automatic prohibited item detection system is thus favorable for speeding up the screening process as well as improving the accuracy of threat detection [3]. Previous studies are primarily based on the bag of visual words model (BoVW) [4]–[6]. Recently, the Convolutional Neural Network (CNN) approaches [7]–[10] have drawn more and more attentions in prohibited item detection. In [11] and [12], Support Vector Machine (SVM) and CNN were combined together for making X-ray baggage image classification and detection. Xu *et al*. [13] proposed an attention-based CNN model of detecting the prohibited items in X-ray images.

Certainly, a big database with enough image samples is indispensable for training the above deep models. Unfortunately, a desirable X-ray image database suitable for training CNN models has not been reported, though GDXray [14], a gray X-ray image database, has been used for performance evaluation in prohibited item detection. Since different materials present different colors during X-ray imaging, a big database of X-ray color images is certainly important for handling the physical problems about prohibited

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

item detection. Usually, it is difficult to collect enough X-ray images containing the prohibited items with great variations in pose and scale. The insufficiency of training images actually results in a great difficulty for developing reliable deep-based approaches suitable for prohibited item detection.

It is undesirable for image database augmentation by translating and rotating the image samples, since no valuable information can be effectively enriched by these ways [15]. In some cases, training a CNN model based on a pre-trained model can slightly improve object detection performance. However, a suitable pre-trained model is not always available for doing this kind of transfer-learning. Recently, the Generative Adversarial Network (GAN) [16] has made a considerable progress in data generation. The achievements of GAN-based methods also show that the images with high quality can be successfully generated indeed [17]–[21]. For instance, Wasserstein GAN with gradient penalty (WGAN-GP) [22], [23], progressive growing GAN (PGGAN) [24], spectral normalization GAN (SNGAN) [25] and self-attention GAN (SAGAN) [26] all could generate the images with high resolution and rich diversity. Hence, GAN-based schemes are encouragingly feasible in data augmentation when the sample count of an database is not enough [15], [27], [28].

FIGURE 1. Some X-ray prohibited item images.

For the task of generating X-ray prohibited item images with high quality and pose variation, training the above GAN models is also difficult since our homemade X-ray image database does not contain enough samples. In addition, the items in baggage are placed randomly and packed tightly, which makes their appearances diverse in imaging. Figure 1 shows some images of handguns, blades and scissors. The prohibited items are random in pose and the backgrounds sometimes clutter during X-ray imaging. This kind of baggage imaging context is unfavorable for GANs to learn some common features of all the objects.

In this paper, we propose a data augmentation method for enriching the X-ray prohibited item images using GAN-based approach. First, the foregrounds containing prohibited items are extracted by a K-Nearest Neighbor (KNN) matting scheme [29] from the collected X-ray security images. Considering the great variations of the prohibited item poses in imaging, the poses of all of the extracted foregrounds are estimated using a space rectangular coordinate system and categorized into 4 or 8 classes. Next, we design an improved GAN model to generate realistic images. Here, the Frechet Inception Distance (FID) [30] is used for the optimal GAN model selection. Then, many new images respectively corresponding to 10 different prohibited item classes are generated using the proposed method. Finally, to verify whether the generated images belong to its corresponding class or not, 10 thousand images with good quality are selected for cross-validation based on a simple CNN model. Thus, the generated images similar to its corresponding original images can be added into the database as new samples. This certainly is helpful for CNN model training.

The rest of this paper is organized as follows. In Section 2, we introduce the collected X-ray prohibited item images and present an image preprocessing method. Section 3 details the proposed GAN model. Section 4 shows some generated images and makes some discussion. In Section 5, we conduct some verification experiments. Finally, Section 6 summarizes this paper.

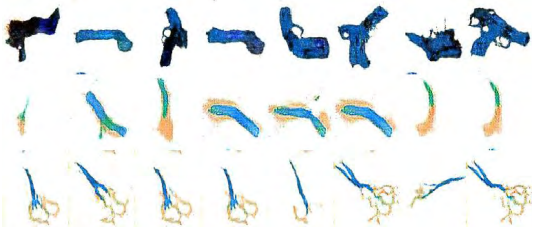


FIGURE 2. Some generated images without preprocessing.

II. IMAGE PREPROCESSING

Most of GAN models for image generation need a large training database, such as ImageNet [31] and LSUN [32]. The insufficiency of training images increase the difficulty of network training. If these images are directly fed into GAN model for unsupervised learning, the network is hard to learn their common features. As shown in Fig. 2, the generated images have unreasonable shapes of handguns, blades and scissors. To solve this problem, we must preprocess the collected X-ray prohibited item images, including foreground extraction and pose categorization.

A. DATABASE

Here, all the X-ray prohibited item images are collected by a X-ray machine working at our Lab. They consist of 10 item classes, such as handguns, lighters, knives etc. Most prohibited item classes contain around 200 images, and some classes contain more than 300 images since they have more item shapes and poses, such as liquid and power bank. All the images are resized to 96×96. Then, we extract prohibited item images using a matting-based method.

B. FOREGROUND EXTRACTING

The prohibited item regions should be segmented firstly from the obtained X-ray images with clutter background. Here, a matting scheme is used to extract foregrounds of X-ray prohibited item images, where the original images and

trimaps are both required. The foregrounds can be extracted by solving Eq. (1),

$$I = \alpha F + (1 - \alpha)B, \tag{1}$$

where I is an original image, F and B respectively denote the foreground and the background of I, and α is a trimap. Details about solving Eq. (1) can be found in [29]. Some important steps for extracting foregrounds of X-ray prohibited item images are shown in Fig. 3. It should be noted that the precision of trimap is important for matting results. Some matting results are shown in Fig. 4.

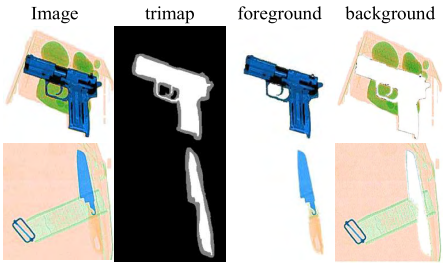


FIGURE 3. Some important steps of extracting the foregrounds from images.

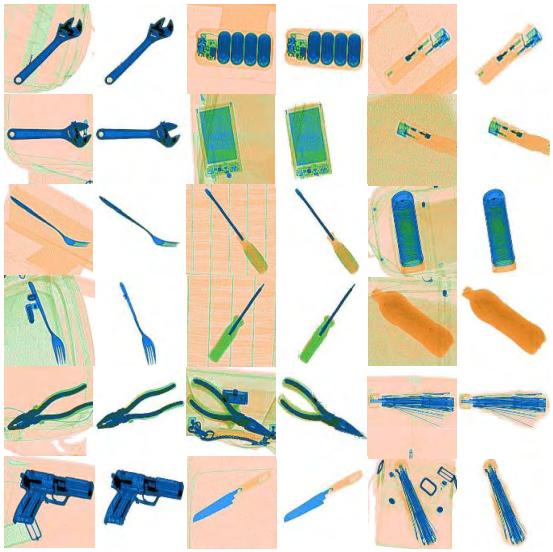


FIGURE 4. Matting results. In the left column, the prohibited items are made of metal materials. The color differences between the foregrounds and the backgrounds are obvious. In the meddle column, the prohibited items contain both metal and organic materials. In the right column, the prohibited items are mainly organic. From the matting results we can see that the foregrounds and the backgrounds can be separated effectively.

C. POSE CATEGORIZING

The poses of the prohibited items vary greatly in X-ray images, which is adverse for GAN model training. To solve the problem, all the poses are estimated by a space rectangular coordinate system and categorized into several classes. Fig. 5 gives an illustration of the space rectangular coordinate system using a handgun as illustration. Different poses corresponding to different angles when handgun rotating along

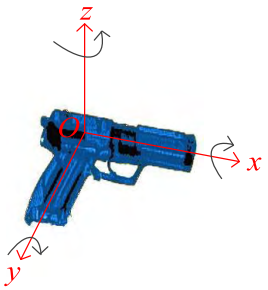


FIGURE 5. A space rectangular coordinate system of handgun.

FIGURE 6. Some categorizing results of different prohibited items.

three axes. Here, we categorize these poses into 4 classes according to the angles that the items rotate along z-axis, respectively $0° \pm 45°$, $90° \pm 45°$, $-90° \pm 45°$ and $180° \pm 45°$. If the items are not axial symmetric, their poses should be categorized into 8 classes. Fig. 6 shows some categorized results of some prohibited items. In this way, the pose varieties in each small classes can be reduced. Next, the X-ray prohibited item images are used for new sample image generation.

III. IMAGE GENERATIVE MODEL

In this section, we introduce an image generative model. The architecture of Generator (G) and Discriminator (D) are constructed based on CNNs. To improve the visual quality of generated images, we use the loss function of consistency term GAN (CT-GAN) [33] and select the appropriate parameters for training. In addition, the FID score is used to evaluate the performance of different GAN models.

A. GAN MODEL

The used GAN model is illustrated in Fig. 7. The basic architecture of G is a deconvolutional neural network. The input is random Gaussian noise vector ($z \sim N(0, 1)$) while the output is a generated image. The basic architecture of D is a convolutional neural network. We focus on adjusting the depth of network and the count of convolution kernels to match the X-ray prohibited item image database better.

Here, the loss function of CT-GAN is used to optimize the model. CT-GAN is proposed based on the improvements of WGAN-GP. Compared with other GANs, it performs better on small databases and is stable in training. It should be mentioned that we make some modifications compared with [33].

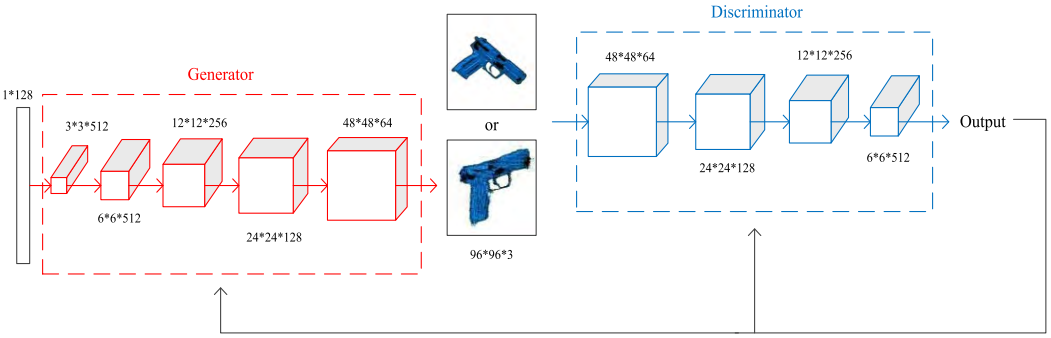


FIGURE 7. The architecture of image generative model.

The loss function is defined as,

$$L = D(G(z)) - D(x) + \lambda_1 GP\mid_{x'} + \lambda_2 CT\mid_{x_1,x_2}, \quad (2)$$

where gradient penalty (GP) and consistency regularization (CT) are defined as follow,

$$GP\mid_{x'} = E_{x'}[(\parallel \nabla_{x'} D(x')\parallel_2 - 1)^2], \quad (3)$$

$$CT\mid_{x_1,x_2} = E_{x\sim P_r}[max(0, d(D(x_1), D(x_2)) - M')], \quad (4)$$

where x' is uniformly sampled from the straight line between the generated data and the real data, x_1 and x_2 are two real images, M' is a constant, λ_1 and λ_2 are two weight coefficients of GP and CT.

Besides the network architecture and loss function, parameters setting is also important for network learning. Batch size, learning rate, weight coefficient, training times, training frequency of D and G are five important parameters. Here, the batch size is 64. The learning rate of D is 0.0002 while the learning rate of G is 0.0003. As for the values of weight coefficients, λ_1 and λ_2 are 0.5 and 5 respectively. (The values given by [33] are not suitable for X-ray prohibited item images.) The generated images can be obtained better when the training frequency of D is the same as that of G. These parameters can be further tuned finely according to the detailed count of training images.

B. FID SCORE

The visual quality of generated images is usually used to test the performance of different GAN models or parameters. But subjective judgment does not work well when the difference is not obviously. Therefore, an objective evaluation metric is necessary. FID score is widely used to evaluate the performance of different models, it performs well in terms of discriminability, robustness and efficiency [34], [35]. Here, FID score is used to compare the performances of GAN models, which is given by Eq. (5),

$$d^2((m_r, C_r), (m_g, C_g)) = \|m_r - m_g\|_2^2 + \mathrm{Tr}(C_r + C_g - 2(C_r C_g)^{1/2}), \quad (5)$$

where d is the FID, the pair (m_r, C_r) corresponds to the real images and the pair (m_g, C_g) corresponds to the generated images. m and C respectively represent means and covariance. FID score shows the differences between the real images and the generated images. The lower score means the better performance.

IV. GENERATED IMAGES AND ANALYSIS

The GAN model is used to generate new X-ray prohibited item images based on the collected real images. First, we compare the images generated by our model with those of other several models. Then, we make a discuss and give some tips for GAN training. All experiments are performed on the computer with a "NVIDIA GeForce GTX 1060 3GB" GPU. The time cost in model training is less than one hour.

A. PERFORMANCE COMPARISON OF DIFFERENT GAN MODELS

Some image samples with different visual quality are generated based on three GAN models (shown in Fig. 8). Firstly, we use the original deep Convolutional Generative Adversarial Network (DCGAN) [18] model to generate images since it is a stable and classical GAN model. However, the generated images are poor in visual quality. There are many noises in images, especially the handgun images. In addition, the shapes of prohibited items are also distorted terribly. Then, the WGAN-GP is employed for new image generation. Although the resolution of most images has been improved, some images still have ghost shapes of prohibited items (the images in red box). The existing models are therefore infeasible to generate X-ray prohibited item images with good quality. Compared with DCGAN and WGAN-GP, the visual quality of images generated by the proposed GAN model have been improved obviously. The object shapes in these images are more realistic.

Apart from the visual quality, we also compare the FID scores of three GANs. The results are listed in Table 1. Our model achieves lower scores, which indicates that this model can approximate the real image distribution better after doing some improvements on network architecture, loss function and parameters setting. Besides, different FID scores are respectively computed on three databases since the difference of item shape complexity. For example, the image distribution

FIGURE 8. Some images generated by different GAN models. The images in the first three rows are generated by DCGAN and the images in the middle three rows are generated by WGAN-GP. The images generated using our model are in the last three rows.

TABLE 1. FID scores of GANs based on three classes of prohibited item images.

Model	Handgun	Fruit Knife	Wrench
DCGAN	230	135	295
WGAN-GP	142	89	244
Ours	86	67	181

of fruit knives is simpler than others. And the quality of fruit knife images is better.

B. THE GENERATED IMAGES OF DIFFERENT PROHIBITED ITEMS

Next, many images are generated using our model based on different prohibited item image databases. Fig. 9 shows some generated image samples. These images have good visual quality of prohibited items, especially the fruit knife images and the images representing liquid materials. Unfortunately, the details of some generated images are still blurred. Besides the visual quality, the diversity of generated images is also important. Compared with other prohibited item images, the images related with liquid materials behave more varieties due the diversity of container appearances.

C. SOME INTERESTING IMAGES

Data augmentation using GAN not only increases sample count, but also enriches the diversity of samples. For instance, some realistic item images with new shapes, colors or poses are generated using the proposed method. These images combine the features of multiple real images. In addition, we also generate some images representing some parts of prohibited items. Actually, there always are some prohibited items that appear partly when X-ray imaging due to occlusion among the contents of baggages. Therefore, these generated images are helpful for enriching the diversity of database. Fig. 10 shows some samples only representing some parts of handguns and fruit knives without handles.

D. TIPS FOR GAN TRAINING

Network architecture, loss function, training process and database have a great influence on GAN training. We introduce the following suggests that are motivated to generate better images:

If the database is small in size, we can add generated images with high quality into database for further model training. Then, the images generated based on the extended database can have better quality. Moreover, the group of the generated images can include more images with new features and less images with ghost shapes.

Apart from the size of database, the object variations in shapes, colors and poses also have impact on model training. Adjusting the learning rate and training times can generate more realistic images. It should be noted that sometimes the quality of images is the by-product of diversity reduction. Therefore, it is necessary in training GANs for making a trade-off between the visual quality and diversity of generated images.

Finally, complex loss functions and model architectures can often improve the performance of GANs. Furthermore, we can make some effective adjustments of models based on special database to generate better images.

V. VERIFICATION

Most generated images are realistic for subjective judgment. But whether or not these images can be recognized correctly by neural networks? In this section, a new X-ray prohibited item image database and the collected real X-ray prohibited item image database are used to perform a cross-validation experiment based on a simple CNN model. Moreover, we extend the comparison experiments of performance evaluation that training with or without generated images.

The new X-ray prohibited item image database consists of 10 thousand generated images. The CNN model here containing three convolutional layers and three full connected layers. The batch size is 64, and the learning rate is 0.0005. First, the real images are used to train the CNN model. After training, the model is applied for the generated image classification. Next, the generated images are used to train the model for doing classification of the real prohibited item images. When training the model, 70% samples of our database are used for training, and the rest samples are used for testing.

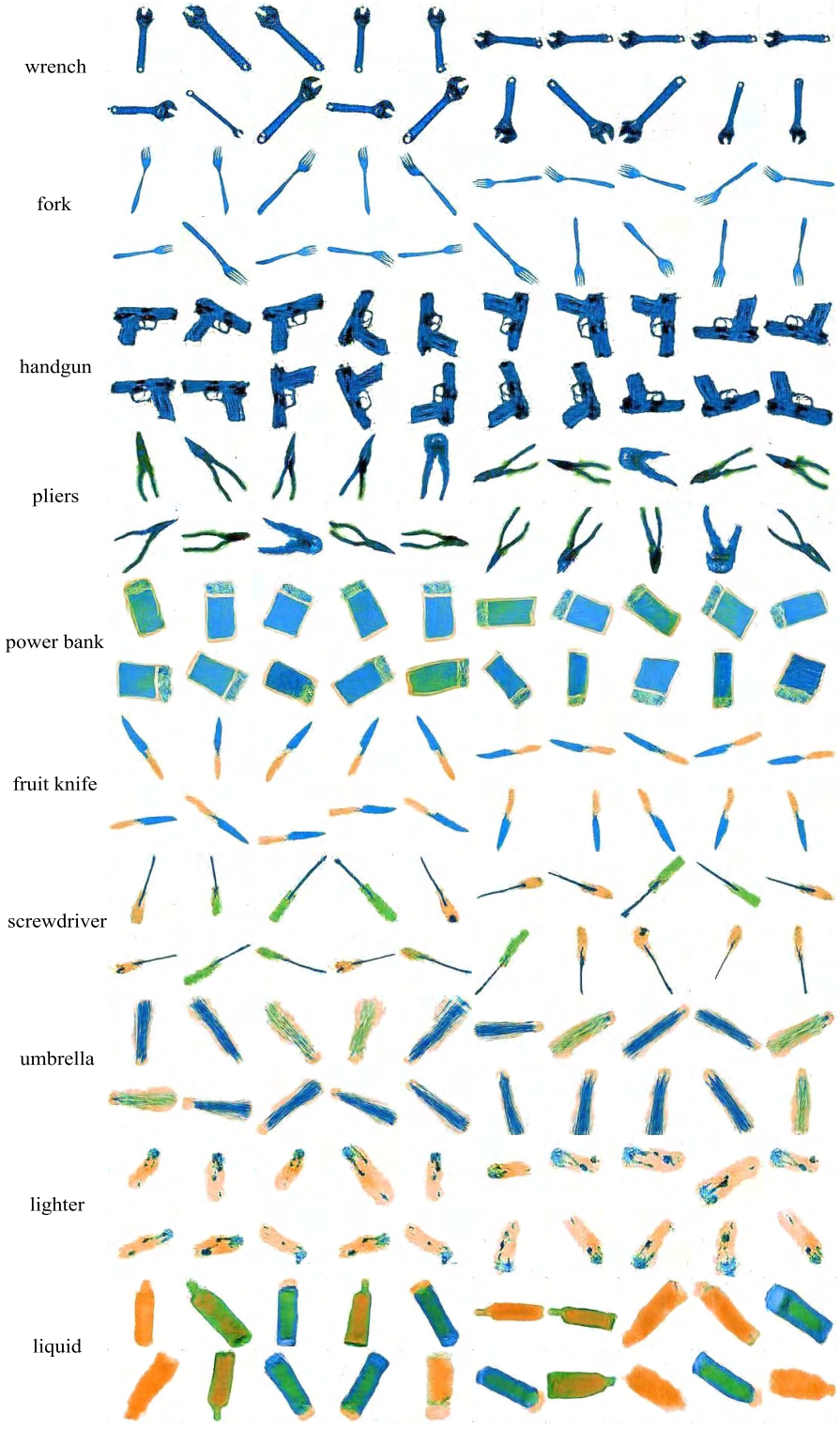


FIGURE 9. Some generated images of different prohibited items.

A. VERIFICATION FOR GENERATED IMAGES

The CNN model is trained for 20 epochs based on real images, and the accuracy of classification is up to 99.16%. Then the trained model is used to classify the generated images which containing 10 classes of prohibited items. Table 2 lists the accuracies of classification over 10 classes of prohibited items. Most of the generated images can be correctly classified into their corresponding classes.

FIGURE 10. Some generated images of parts of knives and handguns.

TABLE 2. Classification accuracy of generated images.

item	accuracy
wrench	99.6%
fork	98.6%
handgun	100%
power bank	99.9%
lighter	100%
pliers	97.5%
fruit knife	98.6%
liquid	100%
umbrella	99.6%
screwdriver	99.6%

FIGURE 11. Some generated images with wrong matching results.

TABLE 3. Classification accuracy of real images.

item	accuracy
wrench	85.23%
fork	89.29%
handgun	89.41%
power bank	94.61%
lighter	89%
pliers	84.88%
fruit knife	92.85%
liquid	97.92%
umbrella	93.69%
screwdriver	94.27%

Some images (shown in Fig. 11) are classified into similar item class. For instance, the forks are wrongly classified into wrenches, the fruit knives are wrongly classified into screwdrivers. This means that the quality of these generated images can be improved furthermore.

B. VERIFICATION FOR REAL IMAGES

Then we use the generated images to train the CNN model. The accuracy of classification is up to 99.93%. The real X-ray prohibited item images are classified by the CNN and the results are listed in Table 3. Unfortunately, the classification accuracy is reduced. The reason is that the diversity of generated images is less than that of real images. Some images of special item poses and positions cannot be generated by GAN model. Fig. 12 gives some image samples.

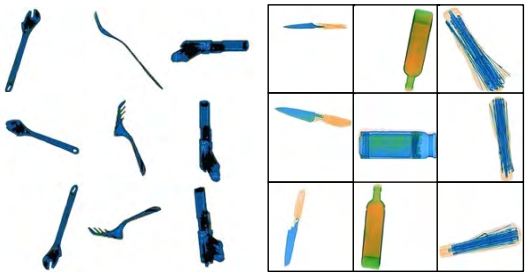


FIGURE 12. Some real X-ray prohibited item images with special item poses and positions. The images of (a) have special poses and the images of (b) have the special positions (not in the center of the images).

The above two experimental results show that most of X-ray prohibited item images generated using the proposed methods can be used as new samples of the database. But the detail information of some generated images is not ideal. In addition, our GAN model could not learn all the distributions of real images. Next, we can generate images using GAN model only based on the images which have specific item poses and positions.

Cross-validation is an effective method to verify whether or not the generated images are suitable for database extension. If we obtain good classification results of both two experiments, all of the generated images are realistic. Furthermore, if the second accuracy is better than the first one, it means that the generated images enrich the diversity of the real images obviously.

C. PERFORMANCE EVALUATION

To further verify that the generated images are useful for model training, we perform a comparison classification experiments. First, 5 thousand real images are used to train the CNN model. This database have been extended using traditional data augmentation methods. Next, 10 thousand generated images and these real images are combined together to train the CNN model. Both two CNN models are verified by a common test database. The test database consists of 3 thousand real X-ray prohibited item images, which are different from the training images.

When the CNN model is trained using the real images, the classification accuracy on test database is 97.80%. If the CNN model is trained using the real images and the generated images, the classification accuracy is up to 98.37%. The better classification accuracy can be obtained when the generated images are used to train the CNN model together with the real images. This shows that the new X-ray prohibited item images generated using the proposed method are desirable for database extending. Here we only generate 1000 images of each item class. If the more generated images are used for model training, we believe that the classification accuracy can be improved furthermore.

VI. CONCLUSIONS

In this paper, a GAN-based method has been proposed to generate images of X-ray prohibited items. After foreground

extracting and pose categorizing, a GAN model was designed to generate many new realistic images with different item poses. The FID score was used to evaluate the performance of GANs. We discussed the quality and diversity of generated images, and gave some tips of handling GAN training in X-ray prohibited item image generation. Finally, a cross-validation experiment verified that most generated images were realistic and suitable for database extending. And we also verified that these generated images were useful for improving the classification accuracy. The generated images can be used to train the models and synthesize the X-ray security images with multiple objects. The synthesized security X-ray images are significant for performance evaluation of prohibited item recognition algorithms. Our work achieved data augmentation for X-ray prohibited item image database effectively.

VII. ACKNOWLEDGMENT

This paper was presented in part at the Chinese Conference on Pattern Recognition and Computer Vision, Guangzhou, 2018. This paper was recommended by the program committee.

REFERENCES

[1] S. Akçay, M. E. Kundegorski, M. Devereux, and T. P. Breckon, "Transfer learning using convolutional neural networks for object classification within X-ray baggage security imagery," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 1057–1061.

[2] D. Mery, E. Svec, M. Arias, V. Riffo, J. M. Saavedra, and S. Banerjee, "Modern computer vision techniques for X-ray testing in baggage inspection," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 47, no. 4, pp. 682–692, Apr. 2017.

[3] D. Turcsany, A. Mouton, and T. P. Breckon, "Improving feature-based object recognition for X-ray baggage security screening using primed visualwords," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Cape Town, South Africa, Feb. 2013, pp. 1140–1145.

[4] D. Mery, E. Svec, and M. Arias, *Object Recognition in Baggage Inspection Using Adaptive Sparse Representations of X-ray Images*. Cham, Switzerland: Springer, 2015, pp. 709–720.

[5] M. E. Kundegorski, S. Akcay, M. Devereux, A. Mouton, and T. P. Breckon, "On using feature descriptors as visual words for object detection within X-ray baggage security screening," in *Proc. Int. Conf. Imag. Crime Detection Prevention (ICDP)*, Madrid, Spain, 2016, pp. 1–6.

[6] Z. Ma, J.-H. Xue, A. Leijon, Z.-H. Tan, Z. Yang, and J. Guo, "Decorrelation of neutral vector variables: Theory and applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 129–143, Jan. 2018.

[7] D. Zhang, D. Meng, and J. Han, "Co-saliency detection via a self-paced multiple-instance learning framework," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 5, pp. 865–878, May 2017.

[8] G. Cheng, P. Zhou, and J. Han, "Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7405–7415, Dec. 2016.

[9] J. Han, D. Zhang, X. Hu, L. Guo, J. Ren, and F. Wu, "Background prior-based salient object detection via deep reconstruction residual," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 8, pp. 1309–1321, Aug. 2015.

[10] Z. Ma, H. Yu, W. Chen, and J. Guo, "Short utterance based speech language identification in intelligent vehicles with time-scale modifications and deep bottleneck features," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 121–128, Jan. 2019.

[11] D. Mery, "Automated detection in complex objects using a tracking algorithm in multiple X-ray views," in *Proc. CVPR WORKSHOPS*, Colorado Springs, CO, USA, May 2011, pp. 41–48.

[12] S. Akcay, M. E. Kundegorski, C. G. Willcocks, and T. P. Breckon, "Using deep convolutional neural network architectures for object classification and detection within X-ray baggage security imagery," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 9, pp. 2203–2215, Sep. 2018.

[13] M. Xu, H. Zhang, and J. Yang, "Prohibited item detection in airport X-Ray security images via attention mechanism based CNN," in *Pattern Recognition and Computer Vision*, J.-H. Lai *et al.*, Eds., Guangzhou, China: Springer, 2018, pp. 429–439.

[14] D. Mery *et al.*, "GDXray: The database of X-ray images for nondestructive testing," *J. Nondestruct. Eval.*, vol. 34, no. 4, p. 42, Dec. 2015.

[15] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification," in *Proc. 15th Int. Symp. Biomed. Imag. (ISBI)*, Washington, DC, USA, Apr. 2018, pp. 289–293.

[16] I. J. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Montreal, Canada, 2014, pp. 2672–2680.

[17] M. Mirza and S. Osindero. (2014). "Conditional generative adversarial nets." [Online]. Available: https://arxiv.org/abs/1411.1784

[18] A. Radford, L. Metz, and S. Chintala. (2015). "Unsupervised representation learning with deep convolutional generative adversarial networks." [Online]. Available: https://arxiv.org/abs/1511.06434

[19] T. Salimans *et al.*, "Improved techniques for training GANs," in *Proc. Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Barcelona, Spain, 2016, pp. 2234–2242.

[20] S. Gurumurthy, R. K. Sarvadevabhatla, and R. V. Babu, "DeLiGAN: Generative adversarial networks for diverse and limited data," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 4941–4949.

[21] Z. Ma, Y. Lai, W. B. Kleijn, Y.-Z. Song, L. Wang, and J. Guo, "Variational Bayesian learning for Dirichlet process mixture of inverted Dirichlet distributions in non-Gaussian image feature modeling," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 449–463, Feb. 2019.

[22] M. Arjovsky, S. Chintala, and L. Bottou. (2017). "Wasserstein GAN." [Online]. Available: https://arxiv.org/abs/1701.07875

[23] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Proc. Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Long Beach, CA, USA, 2017, pp. 5767–5777.

[24] T. Karras, T. Aila, S. Laine, and J. Lehtinen. (2017). "Progressive growing of GANs for improved quality, stability, and variation." [Online]. Available: https://arxiv.org/abs/1710.10196

[25] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida. (2018). "Spectral normalization for generative adversarial networks." [Online]. Available: https://arxiv.org/abs/1802.05957

[26] H. Zhang, I. J. Goodfellow, D. Metaxas, and A. Odena. (2018). "Self-attention generative adversarial networks." [Online]. Available: https://arxiv.org/abs/1805.08318

[27] H. Shi, L. Wang, G. Ding, F. Yang, and X. Li, "Data augmentation with improved generative adversarial networks," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, Beijing, China, Aug. 2018, pp. 73–78.

[28] H.-C. Shin *et al.*, "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent (MICCAI)*, Granada, Spain, 2018, pp. 1–11.

[29] Q. Chen, D. Li, and C. K. Tang, "KNN matting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2175–2188, Sep. 2013.

[30] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local nash equilibrium," in *Proc. Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Long Beach, CA, USA, 2017, pp. 6626–6637.

[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, 2009, pp. 248–255.

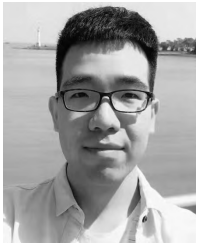
[32] F. Yu, A. Seff, Y. Zhang, S. Song, T. Funkhouser, and J. Xiao. (2015). "LSUN: Construction of a large-scale image dataset using deep learning with humans in the loop." [Online]. Available: https://arxiv.org/abs/1506.03365

[33] X. Wei, B. Gong, Z. Liu, W. Lu, and L. Wang. (2018). "Improving the improved training of Wasserstein GANs: a consistency term and its dual effect." [Online]. Available: https://arxiv.org/abs/1803.01541

[34] Q. Xu *et al.* (2018). "An empirical study on evaluation metrics of generative adversarial networks." [Online]. Available: https://arxiv.org/abs/1806.07755

[35] Z. Ma, A. E. Teschendorff, A. Leijon, Y. Qiao, H. Zhang, and J. Guo, "Variational Bayesian matrix factorization for bounded support data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 4, pp. 876–889, Apr. 2015.

JINFENG YANG was born in 1971. He received the B.Sc. degree from the Zhengzhou University of Light Industry, China, in 1994, the M.Sc. degree from Zhengzhou University, in 2001, and the Ph.D. degree in pattern recognition and intelligent system from the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China, in 2005. He is currently a Professor with the Civil Aviation University of China and Shenzhen Polytechnic. His major research interests include machine learning, pattern recognition, computer vision, multimedia processing, and data mining. He served as a PC Member/ a Referee for international journals and top conferences, including the IEEE TIP, IVC, PR, PRL, ICCV, CVPR, and ECCV. He serves as a Syndic of the Chinese Society of Image and Graphics and the Chinese Association for Artificial Intelligence.



ZIHAO ZHAO was born in 1994. He received the B.S. degree in electronic information engineering and software engineering from Nantong University, in 2016. He is currently pursuing the master's degree with the Civil Aviation University of China. His main research interests include image processing and deep learning.

HAIGANG ZHANG was born in 1989. He received the B.S. degree from the University of Science and Technology Liaoning, in 2012, and the Ph.D. degree from the School of Automation and Electrical Engineering, University of Science and Technology Beijing, in 2018. He is currently with the College of Electronic Information and Automation, Civil Aviation University of China. His research interests include machine learning, image process, and biometrics.

YIHUA SHI was born in 1972. She received the B.Sc. degree from Inner Mongolia Normal University, China, in 1993, and the M.Sc. degree from Zhengzhou University, in 2006. She is currently an Associate Professor with the Civil Aviation University of China. Her major research interests include machine learning, pattern recognition, computer vision, multimedia processing, and data mining.