RESEARCH ARTICLE OPEN ACCESS

Generative Adversarial Network Based Baggage Inspection

Albin Sunny [1], Farzeen Riaz [2], Aman Naveen Sait [3], Farhana Shirin K.N [4], Ajmal E.B [5]

[1], [2], [3], [4] UG Scholar, Dept. Of Computer Science and Engineering, KMEA Engineering College,
[5] Asst. Prof, Dept. Of Computer Science and Engineering, KMEA Engineering College, Kerala - India

ABSTRACT

Convolutional neural networks (CNNs) based methods for automatic discriminant of prohibited items in X-ray images attract attention increasingly. However, it is difficult to train a reliable CNN model using the available X-ray security image databases, since they are not enough in sample quantity and diversity. Recently, generative adversarial network (GAN) has been widely used in image generation and regarded as a power model for data augmentation. In this paper, we talk about different data augmentation methods for X-ray prohibited item images based on GAN. First, the network structure and loss function of the self-attention generative adversarial network (SAGAN) are improved to generate realistic X-ray prohibited item images. Then, the images generated by our model are evaluated using GAN-train and GAN-test.

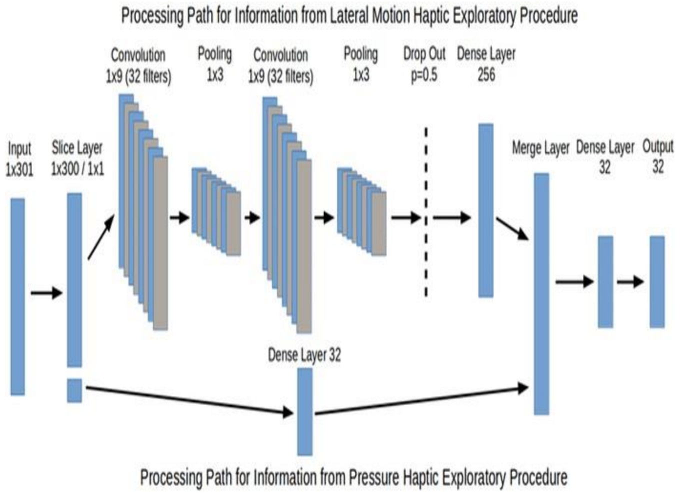
Keywords: GAN, CNN, SAGAN,X-ray security checking image.

I. INTRODUCTION

Artificial intelligence and machine learning have experienced enormous progress in our digital era. The goal of such investigations is to create machines that can identify objects in the same way that humans can. This expansion has presented researchers with a variety of alternatives. In addition, with the introduction of machine learning, deep learning methodologies were introduced, creating a large study field. Artificial intelligence, speech recognition, facial identification, and object detection are just a few of the deep learning applications. For object detection, the convolutional neural network has been widely employed. In comparison to previous classification methods, the amount of pre-processing work required was significantly reduced. Object detection is utilised in a variety of applications, including facial recognition. In this case, the planned For the classification task, CNN is utilised. One of the key technological hurdles in detecting tiny forbidden goods is improving the quality of feature representations. Many academics have worked in recent years to increase the quality of image features using new engines, with feature fusion and learning high-resolution features with broad receptive fields being the most important two categories of approaches. For one reason, because a convolutional neural network (CNN) model is made up of a sequence of convolutional and pooling layers, the semantic information in the deeper layers will be greater. Shallower layers, on the other hand, are less favorable to learning semantics, but they do include more comprehensive information on edges and contours.As a result, combining deep and shallow features in a CNN model helps to increase feature representation quality. Furthermore, because small items take up fewer pixels in pictures, high-resolution features with broad receptive fields retain more characteristics of small objects, which aids in improving small object identification accuracy.

II. CONVOLUTIONAL NEURAL NETWORK:

The convolutional neural network is the neural network used in fields like image recognition and classification. It is a type of neural network that has proven to be effective in image recognition and object detection. CNN has done a good job at recognising faces, objects, and traffic signs. Its success is based on the fact that it is an extension of deep learning algorithms in the convolutional neural network, which does not satisfy the requirements of CNN training in terms of sample quantity and variety. Collecting sample X-ray security checking photos including forbidden goods in various positions and scales by X-ray machine is extremely challenging. In recent years, [6]Generative Adversarial Networks (GANs) have had a lot of success in image generation. There are a lot of GAN models that have been derived. Many derived GAN models have been proposed to improve the quality of the generated images , especially the SAGAN and the BigGAN. Some GAN models are also used for there is not any need of providing high cost and large resources because the standard algorithm can work by applying the relevant filters, the spatial and also the temporal dependencies is also captures by the convolution neural network. We applied in X-ray baggage inspection. Xu et al. [9] proposed a CNN based method with an attention mechanism to detect the X-ray prohibited item images. An et al. [10] built a semantic segmentation net with dual attention to identify the prohibited items. A reliable CNN model for detecting prohibited items needs a perfect X-ray security checking image database, not only in model training but also in model testing. Currently, the available X-ray image database are GDX-ray [11] and SIXray [12].GDX-ray, a grayscale image database, is not suitable to coach model which detect prohibited items in pseudo color image. The SIXray database only includes 8929 images with prohibited items. These databases could not meet the wants of CNN training in sample quantity and variety. The architecture of the convolutional neural network is having the connectivity patterns as that of the neurons in the human brain and was inspired by the visual cortex.



, Fig 1: Convolutional Neural Network

Manual x-ray picture examination, on the other hand, is unreliable and inefficient. Missing detection is a common occurrence in the real world. Passengers wait in line for a long period for a safety check during rush hour, putting security inspection workers under a lot of strain. As a result, a reliable automatic detection system for forbidden objects is required to increase threat detection accuracy while also speeding up the screening process. Convolutional Neural Network (CNN) models have recently demonstrated impressive performance in picture categorization and object recognition. They're also used in the X-ray baggage inspection process. To identify the X-ray forbidden object pictures, researchers Hers was the first to present a CNN-based attention mechanism. To detect the forbidden objects,[4] A created a semantic segmentation net with dual attention. Despite the fact that these strategies increase the efficacy of restricted item identification, they cannot be used in real-world scenarios. Traditional object detectors and deep learning object detectors are the two types of object detectors. With the fast growth of deep learning, particularly the convolutional neural network (CNN), many object identification detectors based on CNN have recently been developed. To detect the items, SSD creates multi-layer feature maps. Adding numerous deconvolution layers to a deconvolutional single shot detector (DSSD) enhances detection accuracy. FSSD concatenates different sizes of feature maps from multiple layers to create a feature pyramid that can directly predict detection outcomes. FSSD concatenates feature maps of various sizes to combine context information. The feature may be fused by concatenating different sizes of feature maps. FSSD is suitable for identifying forbidden objects of big size, but its accuracy in detecting prohibited things of tiny size has to be enhanced. Based on FSSD, a SEM and a Res are introduced to overcome this problem. The viability of methods in data augmentation has been shown. The GAN-based data augmentation approach has recently been applied to supplement the X-ray forbidden item image database. The GAN model is used to create X-ray pictures of items that are forbidden. Their created photographs, on the other hand, only include one restricted object, and the image quality and diversity are not optimal. As a result, we are concentrating our efforts on enhancing the GAN model, which can produce new pictures of higher quality and diversity.

III. GAN STRATEGY

Our simulation approaches can smoothly mix an isolated threat object, like as a pistol, into an X-ray image of baggage, but they can't produce fresh views of the threat object. We simulate fresh danger object samples and reinforce training models using one of the most current GAN architectures to produce more views of threat objects. A typical GAN model consists of a generator and a discriminator network, with the generator attempting to synthesize new danger object pictures from random noise vectors, and the discriminator trained to discern between genuine threat object photos and synthetically created images. [3]The recently suggested progressively increasing regime of GANs (PGGAN) raises the depth of the network by adding convolutional layers after each iteration in the training is finished, which is used to detect danger items in baggage inspection with X-ray pictures using deep learning. More specifically, the model starts with 44 threat object pictures by down sampling the original gallery, then adds new layers to the generator and discriminator and trains with 88 after a few epochs. This iterative multi-resolution training continues until the model is trained for the target picture resolution of 128128 by gradually increasing image resolution and adding extra layers. This gradual growth enhances model convergence while also producing better results at higher resolutions. Separate GAN models were trained for the threat item categories utilized in our experiments: knife, razor blade, gun, and shuriken, where only photos of isolated objects were used to create fresh samples of danger objects. We enrich our gallery by translating the photos in X, Y, and XY dimensions, rotating them, and flipping them in horizontal and vertical directions to impose translation, rotation, and scale invariance in the final detection models. [8]This enhanced data set is then utilised to train our PGGAN model to create new threat object examples that are realistic.

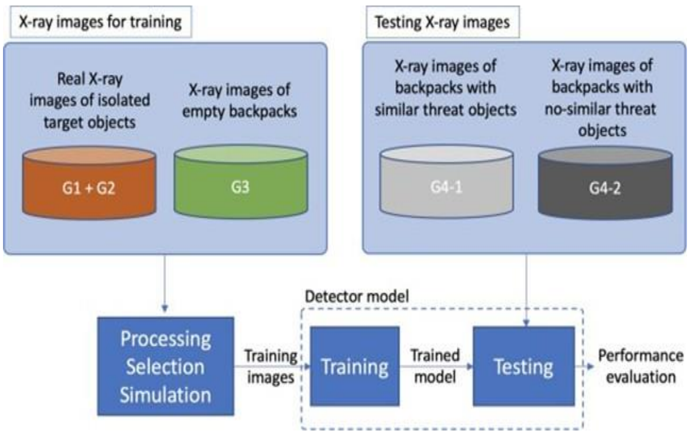


Fig 2:Proposed experimental protocol: In training stage, the X-ray images for training are used. In testing stage, two testing subsets (1 and 2) are used

Fig 3: X-ray prohibited item images without background. The red boxes are the overlapping part of the double prohibited items Different colors represent different materials during X-ray imaging. Blue, green and orange respectively indicate metal,mixture and organic. If the 2 items are overlapping,the color of overlapping dispense with different materials are going to be covered and therefore the color of overlapping give same material will be deepened. as an example, in figure 2, the colour of the overlap between the gun and therefore the pliers is deepened.

IV. DATA SET

The X-ray prohibited item image database has several characteristics which are the important guidance for improving GAN model. 1) The database is small in size. 2) The color of X-ray images provides useful information since different colors represent different materials during X-ray imaging. 3) The different poses of the same prohibited item can be determined by the contour and texture of the images. The contour and texture information of images is crucial to generating X-ray prohibited item images. The color, contour and texture of image are called global image features. To generate high-quality and diverse X-ray prohibited item images, the generation model needs to learn the global image feature well in the small sample database. SAGAN introduces the self-attention mechanism to improve the ability of both Generator and Discriminator to model global structure. The model is widely used to generate realistic images. Because of the small database, SAGAN model can not work well in generating the X-ray prohibited item images. Therefore, we improve the network structure and loss function of the SAGAN model to generate the realistic images. From the section above, we can see that the global images features are very important for X-ray prohibited item images. So improve the network structure to learn the global information of the images on the small database. In order to facilitate the GAN training on the small database, we use the convolution and deconvolution structures as the Discriminator and Generator. [12]We deepen the convolutional network structure so that the convolutional networks can learn the long-range correlation of X-ray prohibited item images. The Discriminator contains six convolutional layers and one fully connected layer. The Generator consists of seven deconvolutional layers and one fully connected layer. We use two sizes of convolution kernels connected in series for Discriminator and Generator. We remove the Batch Normalization of the discriminator layer and only retain the Batch Normalization in the generator. When WGAN-GP and Spectral Normalization are combined, the model is easy to over-fit, so we omit Spectral Normalization in the models, finding that they perform well without it. Compared with WGAN-GP, we find that the hinge loss function is not suitable for the X-ray prohibited item image database. So the WGAN-GP loss function is applied in SAGAN model. Compared to the two-sided penalty used by the original WGAN-GP, we empirically find that one-sided penalty is more suitable for our X-ray prohibited item database, two databases are introduced: X-ray prohibited item image database (Database A) and X-ray security checking image database (Database B). The entire X-ray image databases are manually collected using an X-ray secu- rity inspection equipment working in the laboratory.

IV.1. X-RAY PROHIBITED ITEM IMAGE GENERATION X-RAY PROHIBITED ITEM IMAGE DATABASE

The Database A includes 10 categories of prohibited item images, such as gun, fork, scissor and so on. Each cate- gory involves 200-400 images in 256×256 size, as shown in Figure 3. In order to facilitate the X-ray image synthesis in the subsequent work, we extract the prohibited item fore- ground in images by the method proposed.

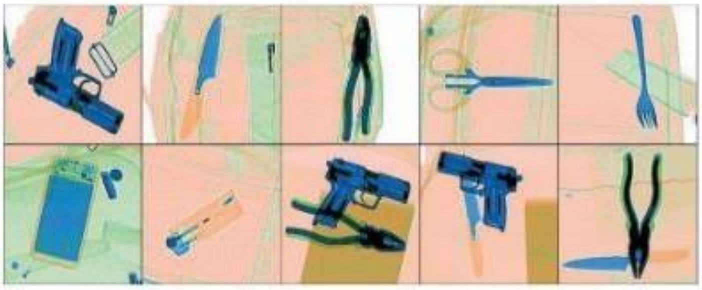


Fig 4. The prohibited item images without background

4.2 X-RAY PROHIBITED ITEM IMAGE GENERATION

The training of most GAN models needs a lot of images. However, the X-ray prohibited item image database is small in size, which would result in model over fitting or model collapse. To generate prohibited item images in high real- ity and rich diversity, we improve the Self-Attention GAN model by matching the prohibited item image database. Then our model is compared with other GAN models based on FID scores. Finally,show some generated images of the proposed mode. Although these generated images are better than others in quality, the edges of the prohibited items are blurred. Therefore, it is difficult for existed models to generate high quality prohibited item images. The last model is the proposed model. Comparing with other GAN models, the visual quality of the generated images using our model improve obviously. However, visual quality is not objective, and we need quantitative comparison. Currently, FID is extensively used for evaluating the generated images. FID has been shown to be more consistent with human evaluation in assessing the realism and variation of the generated samples. The lower FID score, the better model performance. We also compared four models quanti- tatively using FID score. Table were there presents the FID scores of the four models and it can be found that the generated images using our model achieve lowest FID scores. It means that our model can better approximate the distribution of real images than other models. Data augmentation includes sample count and sample diversity. GAN-based method is designed to generate many new prohibited item images. However, the diversity

Of the generated images is restricted by the training database samples. In addition, since it is difficult to obtain many different guns, daggers and other prohibited items in practice, the shape diversity of prohibited items is relatively poor in X-ray security checking images. Here, we propose an image transform method between the natural prohibited item images and the X-ray prohibited item images based on the Cycle GAN to solve these problems.

V. CYCLE GAN-BASED TRANSFORMATIONMETHOD

Cycle GAN is able to achieve image-to-image translation by learning the feature distribution of two domain images, such as color style translation: horse to zebra. In addition, the model is also used for semantic segmentation, edge extraction and so on. According to its powerful ability of image-to-image translation, we realize that the transformation between the natural prohibited item images and the X-ray prohibited item images has the feasibility. By this way, we enrich the shape and pose diversity of the security X-ray prohibited item images. Cycle GAN uses two Generators to learn the mappings of two domains, and two Discriminators to discriminate the generated images and the real images.

5.1 THE TRANSFORMATION RESULT BASED
5.2 ON CYCLE GAN MODEL

The used database includes the natural prohibited item images and the X-ray prohibited item images. Need to collect some real prohibited item images with rich shapes and poses from the internet. The foreground extracting method [is used to extract foreground of the natural images. Next, both the natural prohibited item images and the X-ray prohibited item images are convert to binary images. Compared to the X-ray images, the natural images have richer diversity of the item shape and pose. In this conversion process, we only want to change the color features of the prohibited item images, retaining the shape and pose features. So, the X-ray prohibited item images and their corresponding binary images are used to train the Cycle GAN model. After training, the model could achieve the transformation between the natural style and the X-ray style. Then, the cor- responding binary images of the natural images could be transformed into X-ray images by the model. Figure shows the transformation results of some hand- guns and hammers from natural images to X-ray images.

Fig 5: Some Generated Image Samples

Some transformation results between the natural images and the X-ray images has been shown here.

Each column contains two sets of images, which are a natural image, a binary image, and an X-ray image from left to right. This method can only convert the color distribution of the items while ensuring that the shape and pose of the prohibited items in images are unchanged. Take the handgun as an example. Since the handgun is usually made of metal and the main color under X-ray is blue, the color of the generated handgun image is also blue and its special texture is also not changed. Therefore, the rich features of the natural item images are transformed into the X-ray images. In this Section, a method based on the Cycle GAN model is proposed to convert the natural images to the X-ray images. The images generated by this method can effectively enrich the variety of the shapes and pose of the prohibited item images, especially the prohibited items which are difficult to obtain and have a wide variety of shapes and pose.

VI. .X-RAY SECURITY CHECKING IMAGE SYNTHESIS

In previous Sections, we mainly introduce the data augmenta- tion method of X-ray prohibited item images. Through these methods, we generated many prohibited item images with high quality and rich diversity. In order to get X-ray security checking images, these prohibited item images are synthe- sized with background images. The background images are also collected manually in our Lab. Thus, in this Section, a method for X-ray security checking images synthesizing is introduced.

Different materials present different colors under X-ray. On the one hand, when the prohibited items of multiple materials are placed overlaid, X-ray imaging of high-priority materials covers with low-priority images. On the other hand, the color of the overlapping prohibited items with the same material is deepened in the X-ray imaging. Thus, according to these characteristics of X-ray imaging, we choose the following methods to synthesize the X-ray security checking images. These images are used to synthesize X-ray security checking images. In the process of synthesis, in order to make the composite images match the priority of X-ray image imaging, we first synthesize the prohibited item of organic material, then the mixture and finally the metal items. We also combine the different prohibited item images with the background images randomly, such as item class and item position. In addition, the size of the prohibited items varies randomly according to their actual size. By this way, the diversity of synthesized X-ray security checking images is enriched. In the previous Sections, we not only talk about the generation of single-item images but also generate overlapping multi-item images. To verify whether the expanded database has the effect of data augmentation, there is desire of designing a comparison experiment of performance evaluation by training object detection models with or without synthesized images. If the performance of object detection model trained by the enlarged database has improved, it will show that our data augmentation method works. In this experiment, apply a classical SSD model as the object detection model. First, we introduce the database in the experiment. The original database is the Database B. These images are collected manually, and there are only 4500 images. We synthesize4200 X-ray security checking images containing 7 types of prohibited items, which is same with the category of the real images. The size of the synthesized images is also 512×512. The enlarged database (named Database C) consists of the real X-ray security checking images and the synthesized X-ray security checking images. Thus, the Database C includes 8700 X-ray security checking images with seven classes. Based on the Database B and Database C, we manually add the bounding-box for each prohibited item. We randomly divide the Database B into two subsets for training and testing. The Database C is also divided into two subsets for training and testing.

From experimental results we find that the accuracy of the SSD model trained by the Database C is higher than the SSD model trained by the Database B, which the mAP increased by 5.6%. Since the large size and different colors of some prohibited items such as power bank and pliers, it is easy to be detected by the SSD model. Thus, the improvement of detec- tion model performance after data augmentation is limited. However, the AP of other prohibited items like lighter, fork and knife has been greatly improved after data augmentation. This shows that data augmentation based on our method can help the detection model detect these prohibited items with small size and similar colors better, especially the AP of lighter is improved from 63.9% to 83.4%. However, the AP of gun declined. We speculate that the capacity of the SSD model is limited so that the model can not learn the feature of the real images from the enlarged database very well. Here we only synthesized 4200 X-ray security checking images. If there are more synthesized images for SSD model training, we think the performance of SSD model can be fur- ther improved. Therefore, it is concluded that the synthesized images based on our method can enhance the X-ray security checking image database.

CONCLUSION

So the proposed work has the convolutional neural network with transfer learning and the majority voting scheme works well in comparison with the actual work in the transfer learning based object detection using CNN.The proposed work gives fine granularity to work done earlier.

In the next one , we adopt the FSSD to detect the prohibited items in X- ray security image checking. We find that the synthesized X-ray security checking images are very close to the real images. The high accuracy of GAN-train indicates that the quality and diversity of the generated images is similar to real images. The GAN-test with a high value denotes that the generated images are a realistic approximation of the distribution of real images. Security inspection images and introduce semantic enrichment module and residual module to improve the detection accuracy of small prohibited items. From the result, we can conclude that these two modules are beneficial for improving detection accuracy of prohibited items with small size. The research is meant to advance the study of detection of threat objects in X-rays. Through the proposed methodology, emphasis is placed on the ease in creating simulated images and their almost complete independence from the test images. In this way, a solid alternative is generated for the training of new deep learning models, without the need to have access to more complex databases and difficult public concession. Together with the proposal of a clear evaluation protocol that can be replicated, and therefore compared to future research that wishes to corroborate its methodology. The idea of our work is to show how to use known object detection strategies in baggage inspection, rather than to develop new models from scratch. However, as we can see from our experiments, these strategies cannot be implemented in a straightforward manner, because they are not effective when the number of available X-ray images for training is low. Unfortunately, the databases in X-ray testing are rather limited. To overcome this problem, propose a strategy for deep learning training that is performed with a low number of target-free X-ray images with superimposition of many simulated targets. The simulation is based on absorption's law of X-rays that allows us to superimpose different layers. Using this method, it is very simple to generate additional training data. Thus by the experiments done and this particular field of xray image security inspection could be efficiently lead by deep learning and frameworks like GAN etc. further research and advancement can be introduced for implement also acceptable.

REFERENCES

[1] Akcay S, Kundegorski M E, Willcocks C G and Breckon Y P,
IEEE Transactions on Information Forensics and Security 13, 2203 (2018).

[2] Xu Mao-shu, Zhang Hai-gang and Yang Jin-feng, Prohibited Item Detection in Airport X-Ray Security Images via Attention Mechanism Based CNN, Chinese Conference on Pattern Recognition and Computer Vision, 429 (2018).

[3] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, ''Going deeper with convolutions,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Boston, MA, USA, Jun. 2015, pp. 1–9.

[4] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[5] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, ''Rethinking the inception architecture for computer vision,'' in Proc. IEEE Conf.Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 2818–2826.

[6] R. Girshick, ''Fast R-CNN,'' in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Santiago, Chile, Dec. 2015, pp. 1440–1448.

[7] 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. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 779–788.

[8] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, ''SSD: Single shot multibox detector,'' in Proc. Eur. Conf. Comput. Vis. (ECCV), Amsterdam, The Netherlands, Oct. 2016, pp. 21–27.

[9] 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.

[10] J. An, H. Zhang, Y. Zhu, and J. Yang, ''Semantic segmentation for prohib- ited items in baggage inspection,'' in Proc. Int. Conf. Intell. Sci. Big Data Eng. (ISCIDE), Nanjing, China, Oct. 2019, pp. 495–505.

[11] D. Mery, V. Riffo, U. Zscherpel, G. Mondragón, I. Lillo, I.Zuccar,H. Lobel, and M. Carrasco, ''GDXray: The database of X-ray images for nondestructive testing,'' J. Nondestruct. Eval., vol. 34, no. 4, p. 42, Dec. 2015.

[12] C. Miao, L. Xie, F. Wan, C. Su, H. Liu, J. Jiao, and Q. Ye, ''SIXray: A large- scale security inspection X-ray benchmark for prohibited item discovery in overlapping images,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Long Beach, CA, USA, Jun. 2019, pp. 2119–2128.

[13] I. Goodfellow, J. Pouget-Abadie, and M. Mirza, ''Generative adversarial nets,'' in Proc. Int. Conf. Neural Inf. Process. Syst. (NIPS), Montreal, QC, Canada, Dec. 2014, pp. 2672–2680.

[14]A. Radford, L. Metz, and S. Chintala, ''Unsupervised representation learning with deep convolutional generative adversarial networks,'' 2015, arXiv:1511.06434. [Online].Available: http://arxiv.org/abs/1511.06434

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

[16] H. Zhang, I. Goodfellow, D. Metaxas, and A.

Odena, ''Self-attention gener- ative adversarial networks,'' 2019, arXiv:1805.08318.[Online]Available:https://arxiv.org/abs/1805.0 8318

[17] A. Brock, J. Donahue, and K. Simonyan, ''Large scale GAN training for high fidelity natural image synthesis,'' 2018, arXiv:1809.11096. [Online]. Available: http://arxiv.org/abs/1809.11096

[18] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, ''Image-to-image translation with conditional adversarial networks,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, Jul. 2017, pp. 1125–1134.

[19] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, ''Unpaired image-to-image translation using cycle-consistent adversarial networks,'' in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Venice, Italy, Oct. 2017, pp. 2223–2232.