Automated Threat Detection In X-Ray Imagery For Advanced Security Applications

Amit Upreti
Department of Computing Science
University of Alberta
116 St 85 Ave, Edmonton, AB T6G 2R3
Email: upreti@ualberta.ca

Rajat Bothra Jain
Department of Computing Science
University of Alberta
116 St 85 Ave, Edmonton, AB T6G 2R3
Email: bothraja@ualberta.ca

***Abstract*—In areas of high security, like airports, etc, X-Ray machines are used to scan baggage to look for hazardous objects such as guns, knives, razor blades. But, in a manual scan, it is easy to miss some details. Also manual scan is a tedious and time-consuming process. We investigate and compare several algorithms for detection of harmful/hazardous objects like razor blades/handguns etc in X-Ray images of travellers' baggage. Deep Convolutional neural networks such as RCNN, Detectron, RetinaNet and Yolo has shown great results in object detection and recognition. We plan to use the object detection techniques and apply them to improvise upon the already existing Automatic baggage screening methods.**

***Index Terms*—Baggage screening, Deep Learning, Convolutional Neural Networks, Object Detection Algorithms, X-ray Images**

1. Introduction

Luggage screening is a very important part of the airport security risk assessment and clearance. Identifying and detecting dangerous objects and threats in baggage carried on board aircraft plays an important role in ensuring and guaranteeing passengers security and Safety [9].

In addition to human error due to workload and fatigue, a wide range of shapes and rotation of hazardous items make it difficult to recognize them. Due to the complex nature of the task, the literature suggests that human expert detection performance is only about 80-90% accurate [8]. Deep convolutional neural networks have already shown good results as demonstrated by Zhao et al [12]. Our goal is to create a model that can analyze an X-Ray image and detect hazardous objects. Analyze the model performance using the standard metrics like accuracy, mean average precision (mAP) etc. Try to improve the performance of automatic baggage screening by exploring the latest object detection techniques.We are also aiming to use Data augmentation methods as discussed by [3] and Threat Image Projection by [4].

2. Literature Review

In 2011, Bastan et al. [2] proposed the Bag-of-Visual-Words on Baggage X-Ray Images and concluded that straight forward application of BoW on X-ray Images doesnot perform well as it does on regular images. 2013, Turcsany et al [11] proposed classification technique for object detection in X-ray baggage imagery using primed visual words in an SVM classifier framework. Primed visual words are obtained through class-specific clustering of feature descriptors and used to encode images in bag-of-words model. This differs from the traditional approach, which combines the feature set of positive and negative classes during the clustering process when generating a codebook. Modification to the clustering stage of the traditional bag-of-words framework creates an image representation scheme that further facilitates the separation of positive and negative class.This method significantly outperforms the previous work of Bastan et al. [2]. People start using CNN networks for Object Detection in X-ray images.Akcay et al [1] examines the applicability of traditional sliding window convolutional neural network (CNN) detection and the relative performance of contemporary object detection strategies for region based object detection techniques – Faster Region-based CNN (R-CNN) and Region Based Fully Convolution Networks(R-FCN) on X-ray securities images. Use Transfer learning due to limitation of training dataset of 11,627 samples (5,867 training, 2,880 validation and 2,880 test samples). The Faster RCNN and R-FCNN provide superior results than traditional sliding window driven CNN(SW-CNN) approach.Faster RCNN with VGG16, pretrained on the ImageNet dataset, achieved 88.3 mAP for a six object - firearm, firearm-components, knives, ceramic knives, camera and laptop detection in X-ray dataset. R-FCN with ResNet-101, yields 96.3 mAP for the two class firearm detection problem and requires 100 milli second computation per image.

Petrozziello et al uses a thresholding algorithm followed by normalization to preprocess the images which isloates the dense materials such as metals which can help to reduce the benign background information within the image [9]. It also demonstrated that CNN perform better than auto encoders

Model	Network	mAP	camera	laptop	gun	gun component	knife	ceramic knife
SWCNN	AlexNet	60.8	68.2	60.9	74.8	71.4	21.2	68.3
	VGGM	63.4	70.7	63.7	76.3	73.1	24.6	71.9
	VGG16	64.9	70.1	72.4	75.2	75.7	22.3	73.4
	ResNet-50	67.1	69.2	80.1	74.7	76.1	31.4	71.3
	ResNet-101	77.6	88.1	90.2	83.1	84.8	39.2	80.3
RCNN	AlexNet	64.7	79.1	81.5	85.3	58.2	18.8	65.8
	VGGM	68.6	79.9	85.5	86.9	65.8	21.0	72.3
	VGG16	77.9	88.8	95.4	87.6	83.2	30.4	81.9
FRCNN	AlexNet	78.8	89.3	75.6	91.4	87.4	46.7	82.3
	VGGM	82.3	90.0	83.4	91.8	87.5	54.2	86.9
	VGG16	**88.3**	88.1	91.8	92.7	93.8	72.1	91.2
	ResNet-50	85.1	84.4	87.9	91.6	90.1	67.7	88.9
	ResNet-101	87.4	85.7	90.4	93.1	91.1	73.2	90.7
RFCN	ResNet-50	84.6	89.4	92.8	93.2	91.8	50.6	89.6
	ResNet-101	85.6	88.7	90.6	94.2	92.5	55.6	92.0

Figure 1. Detection results of SW-CNN, Fast-RCNN (RCNN) [28], Faster RCNN (FRCNN) [20] and R-FCN [21] for multi-class problem (300 region proposals).

for baggage screening.

Bhowmik et al. [3] discusses on how to synthetically create x-ray images with prohibited items(Data augmentation). This creates an opportunity for as the existing public domain datasets such as GDXray [6] contains lesser clutter ,overlap and contains limited categories. SIXray [7] has dataset with clutter but has smaller number of categories.

2020, Hassan et.al [5] proposed a novel method to overcome the retraining requirement of framework across multiple scanner-specification. It uses meta-transfer learning-driven tensor shot detector that decomposes the candidate scan into dual-energy tensors and employs a meta-one-shot classification backbone to recognize and localize the cluttered baggage threats. This method can be well generalized for multiple scanner specifications due to its capacity to generate object proposals from the unified tensor maps rather than diversified raw scans. On the SIXray dataset, the proposed framework achieved a mean average precision (mAP) of 0.6457, and on the GDXray dataset, it achieved the precision and F1 score of 0.9441 and 0.9598, respectively.

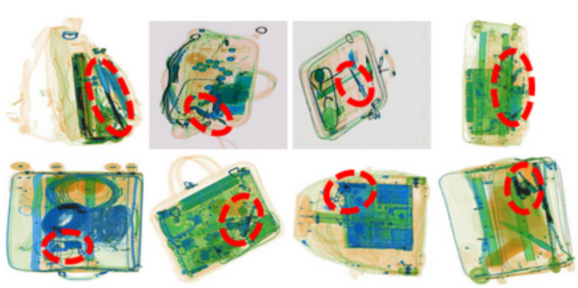

Figure 2. Sample Dataset (SIXRay)

3. Methodology

3.1. Synthetic Image generation by Threat Image Projection

The GDXray dataset has limited number of images with less diversity. Most of the images are from few selected baggages which are taken from various angles. We explored the TIP approach as proposed by Bhomik et al [3] to introduce diversity in the dataset. It involved 3 steps

1) Threat signature Transformation
2) Insertion Position Determination
3) Image Compositing

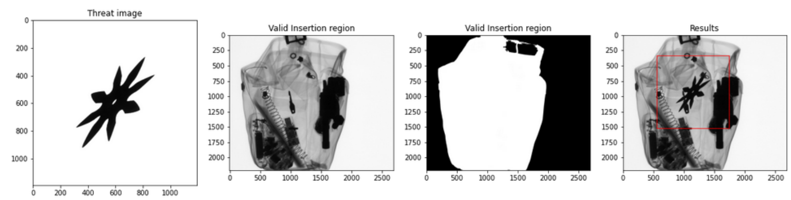

Figure 3. Synthetic Image Generation

We generated a total of 2000 synthetic Xray baggage images (600 for each class Shuriken, Razor Blade, Knife and 200 for Guns) for the following classes :

- Guns
- Shuriken
- Razor blade
- Knife

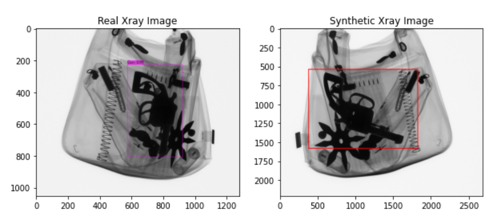

Figure 4. Comparision of Real Image VS Synthetic Image

3.2. Detection Strategies

We used Yolov3-tiny [10] for the purpose of our classification. We created two models one from Real Dataset from GDXRay [300 images] and another model from our Synthetic Image Dataset [2000 images]. We used Yolov3-tiny because of it performance, so that we can evaluate the results faster. Instead of training the model from scratch we have used pre-trained Yolov3-tiny weights for training.

Accuracy Comparison

When training Real Image Dataset we separated 210 data for training and 90 data for testing. Similarly for Synthetic dataset we separated 1600 data for training and 400 data for

Set		Baggage Data
Training	Series	B0046
	Images	1-200
	Series	B0047
		1-10
Testing	Series	B0047
	Images	11-90

Figure 5. Images of GDXRay Used In Our Experiments

Classes	Testing Accuracy	
	Real Model(on Real Dataset)	Synthetic Modell(on Real Dataset)
Guns	93.59	79.26
Knives	94.82	0
Razor	55.05	0.34
Shuriken	86.59	12.37

Figure 6. Test Real Dataset

testing, maintaining the diversity in both training and testing datasets.

We tested Both of our models first on their individual testing data and then tested on each other.

Testing on real dataset : yolov3-tiny trained on real data performed well in classifying objects, however for smaller objects such as razor/blade accuracy is quite low. We are able to get good accuracy for guns and extremely low accuracies for others using yolov3-tiny model trained on synthetic data.

yolov3-tiny, uses low resolution feature and sometimes object features get too small to be detectable may be reason for low accuracy of small objects on real dataset.

Testing on synthetic dataset : yolov3-tiny trained on synthetic data performed well in classifying all the objects. However, while testing on synthetic data on model trained with real data we have received extremely low accuracy .This highlights the difference in features between synthetic and real dataset created.

4. Discussion

We explored the use of Image Projection to increase the diversity in data and to provide robustness to model when training. These techniques can be very helpful in Deep Learning especially in cases of baggage screening where there is lack of Labelled publicly available data. Popular Datasets available For Training mostly contains only 4 classes of Objects: Knife, Razor/Blade, Shuriken and Gun. But actual machines used at site need to identify more than 12 classes of objects : Gun, Knife, Bomb, Scissor, Blade, Spanner, Tools, GunParts, Mobile, BatteryBank, Charger, Battery etc. Image Projection as seen in our observations offers promising result for gun detection . This technique can be used to train model on objects whose datasets are not easily available as our data generation technique becomes

Classes	Testing Accuracy	
	Real Model(On Synthetic Dataset)	Synthetic Model(On Synthetic Dataset)
Guns	5.37	100
Knives	0.0	98.47
Razor	0.06	90.21
Shuriken	1.88	98.52

Figure 7. Test Synthetic Dataset

closer to real data our dependency on new dataset decreases drastically without impacting the accuracy .

During creating synthetic images, we had a limitation where we had a single background/target image and few threat images per classes. Both of them were simply taken from various angle. Although we were able to diversity by increasing the dataset, but it might not be enough for a system as there is little variation in the images that were used to generate synthetic images.

5. Conclusion and Future Work

Image Projection is an efficient method to introduce diversity and boost the generalization of Deep learning Models. We were able to create a large number of automatically labelled images with minimal resources. Such techniques have can save countless hours of human work which is needed during labelling of dataset. There is a challenge to blend images into occluded backgrounds(where there are multiple overlaps of objects) where an unnatural projection is done. Our Training Results shows we are able to get reasonable accuracy for guns by training the model on synthetic data. Here we used GDXray which had very limited number of sample images, We believe using a more diverse dataset such as SIXray can boost the Threat Image Projection even further.

References

[1] Samet Akçay, Mikolaj E Kundegorski, Michael Devereux, and Toby P Breckon. Transfer learning using convolutional neural networks for object classification within x-ray baggage security imagery. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 1057–1061. IEEE, 2016.

[2] Muhammet Baştan, Mohammad Reza Yousefi, and Thomas M Breuel. Visual words on baggage x-ray images. In *International Conference on Computer Analysis of Images and Patterns*, pages 360–368. Springer, 2011.

[3] Neelanjan Bhowmik, Qian Wang, Yona Falinie A Gaus, Marcin Szarek, and Toby P Breckon. The good, the bad and the ugly: Evaluating convolutional neural networks for prohibited item detection using real and synthetically composited x-ray imagery. *arXiv preprint arXiv:1909.11508*, 2019.

[4] Victoria Cutler and Susan Paddock. Use of threat image projection (tip) to enhance security performance. In *43rd Annual 2009 International Carnahan Conference on Security Technology*, pages 46–51. IEEE, 2009.

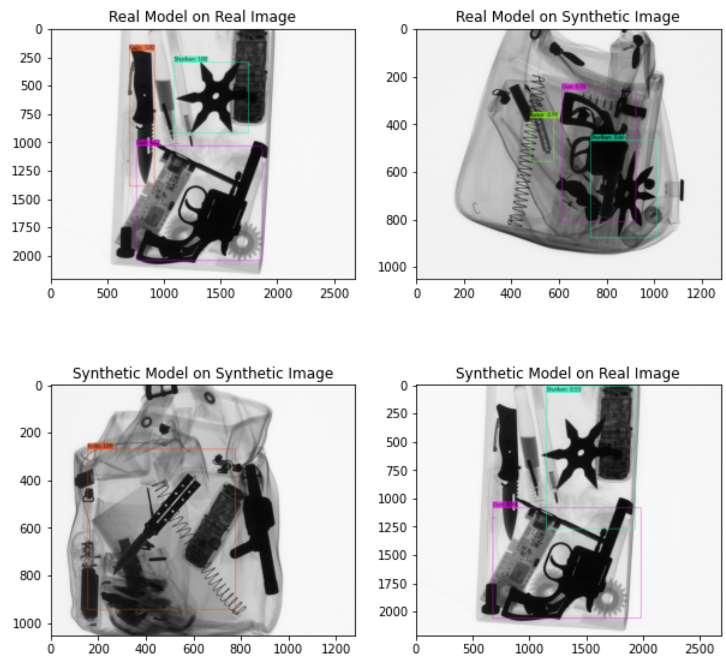

Figure 8. Modelling results on Real and Synthetic

[5] Taimur Hassan, Muhammad Shafay, Samet Akçay, Salman Khan, Mohammed Bennamoun, Ernesto Damiani, and Naoufel Werghi. Meta-transfer learning driven tensor-shot detector for the autonomous localization and recognition of concealed baggage threats. *Sensors*, 20(22):6450, 2020.

[6] Domingo Mery, Vladimir Riffo, Uwe Zscherpel, German Mondragón, Iván Lillo, Irene Zuccar, Hans Lobel, and Miguel Carrasco. Gdxray: The database of x-ray images for nondestructive testing. *Journal of Nondestructive Evaluation*, 34(4):1–12, 2015.

[7] Caijing Miao, Lingxi Xie, Fang Wan, Chi Su, Hongye Liu, Jianbin Jiao, and Qixiang Ye. Sixray: A large-scale security inspection x-ray benchmark for prohibited item discovery in overlapping images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2119–2128, 2019.

[8] Stefan Michel, Saskia M Koller, Jaap C de Ruiter, Robert Moerland, Maarten Hogervorst, and Adrian Schwaninger. Computer-based training increases efficiency in x-ray image interpretation by aviation security screeners. In *2007 41st Annual IEEE international Carnahan conference on security technology*, pages 201–206. IEEE, 2007.

[9] Alessio Petrozziello and Ivan Jordanov. Automated deep learning for threat detection in luggage from x-ray images. In *International Symposium on Experimental Algorithms*, pages 505–512. Springer, 2019.

[10] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[11] Diana Turcsany, Andre Mouton, and Toby P Breckon. Improving feature-based object recognition for x-ray baggage security screening using primed visualwords. In *2013 IEEE International conference on industrial technology (ICIT)*, pages 1140–1145. IEEE, 2013.

[12] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu. Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11):3212–3232, 2019.