Using Deep Neural Networks to Address the Evolving Challenges of Concealed Threat Detection within Complex Electronic Items

Neelanjan Bhowmik[1], Yona Falinie A. Gaus[1], Toby P. Breckon[1,2]
Department of {Computer Science[1] | Engineering[2]}, Durham University, UK

I. Introduction

'We identified vulnerabilities with TSA's screener performance, screening equipment, and associated procedures'[1] - report published by the Office of Inspector General (OIG), U.S. Department of Homeland Security on covert testing of Transportation Security Administration (TSA) screening checkpoint effectiveness (2017). Undercover investigators managed to conceal threat items, such as explosives, imitation guns, knives, etc. through the security checkpoints 70%∼80% of the time at varius US airports[2]. Although these statistics show an evident improvement compared to two years previously, the failure rate above 70% is cause for concern.

'Aviation is growing, and that is generating huge benefits for the world. A doubling of air passengers in the next 20 years could support 100 million jobs globally.' - Alexandre de Juniac (Director General and CEO, IATA). In 2017 it is estimated by International Air Transport Association (IATA) that commercial airlines carried approximately 4.1 billion passengers worldwide, which is an increment of 7.3% compared to 2016[5]. As illustrated in Fig. 1, the number of air travelers has grown tremendously over the last decade from ∼2 billion in 2004 to ∼4.6 billion in 2019. With this trend, it is estimated that this volume will reach ∼8.2 billion in 2037. Furthermore, it is estimated that between the year 2016 and 2040, air cargo traffic will increase by 2.5% with a 1.9% increment in aircraft movements[6] (as depicted in the Fig. 2).

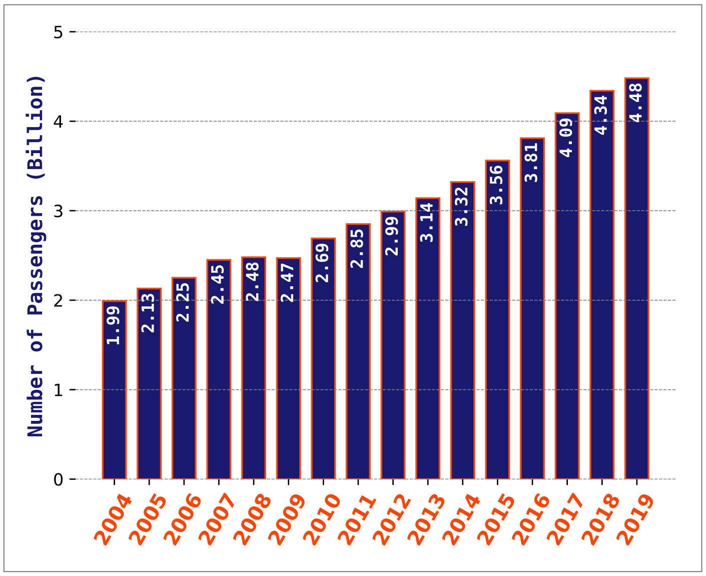

Fig. 1. Statistics on a number of scheduled passengers handled by the airline industry from 2004 to 2019[4].

[1] https://www.oig.dhs.gov/sites/default/files/assets/2017/OIG-17-112-Sep17.pdf

[2] https://abcnews.go.com/US/tsa-fails-tests-latest-undercover-operation-us-airports/story?id=51022188

[4] https://www.iata.org/pressroom/facts_figures/fact_sheets/Documents/fact-sheet-industry-facts.pdf

[5] https://www.iata.org/pressroom/pr/Pages/2018-09-06-01.aspx

[6] https://aci.aero/news/aci-world-report

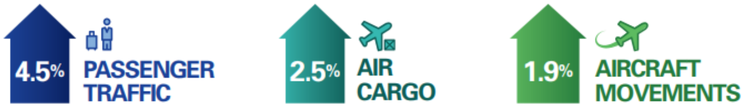

Fig. 2. Global average annual growth forecast 2016-2014.

Indeed, with this increasing volume of traffic, we need to ensure a proficient future system for aviation securing capable of addressing the evolving threat landscape that emanates from broader global geo-political events. In step with this, the complexity of aviation security task itself evolves with air passengers increasingly carrying a wide variety of electronic and electrical items within their baggage. Currently multiple-view X-ray baggage security screening is widely used to maintain aviation and transport including the screening of these electronics items. To address these future challenges of increasing volumes and complexities, the recent focus on the use of automated screening approaches is of particular interest. This includes the potential for automatic threat detection as a methodology for concealment detection within complex electronics and electrical items screened using low-cost, 2D X-ray imagery (single or multiple views).

At this point, we can ask - what is the potential for such a device borne threat, which could be hidden in an electronics device, such as a laptop? In February 2016, a Daallo Airlines aircraft was damaged by mid-air laptop improvised explosive device (IED) explosion[7]. Whilst in March 2016 a laptop IED exploded at the security checkpoint in Mogadishu, Somalia prior to flight. Ultimately, we readily find that the threat is real and evident.

Passenger baggage is currently inspected manually using dual-energy multiple-view X-ray imaging. Furthermore, several human factors such as stress (during peak hours), tediousness (volume of items to be screening) or uncertainty (variety of item) can cause human operators to respond differently. The threat concealment can also be also very subtle and very well hidden within the electronics item, making it challenging for a human operator to identify. For example, as depicted in Fig. 3, the threat portion are camouflaged with other laptop parts. With both increased passenger throughput in the global travel network and an increasing focus on wider aspects of extended border security (e.g. freight, shipping, postal), this poses both a challenging and timely automated image classification task.

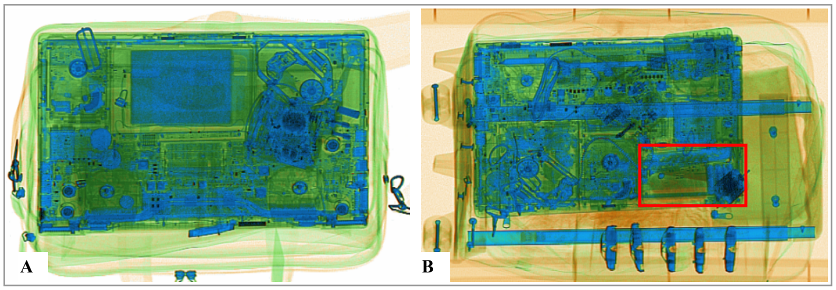

Fig. 3. Spot the difference: Exemplar consumer electronics items within X-ray security imagery with concealed threat region highlighted (red box, B) while other is a benign item (A).

Considering the key challenge of identifying subtle threats within the X-ray imagery, we examine three different strategies to illustrate an automated pipeline for, *firstly:* considering the full X-ray image containing the object without prior object localization, *secondly:* the object level segmentation of such items - *focus on locating the item within the image first, then determine if it contains a threat or not?* and *thirdly:* the object over-segmentation - *is this part of this complex item within the image anomalous?* Automatically segmenting varying electronic items within X-ray security images such that individual objects, and their sub-components, can be isolated within the X-ray image to facilitate improved screening both by human operators and automated screening algorithms enables the detection of anomalous sub-

[7] https://www.bbc.co.uk/news/world-africa-35521646

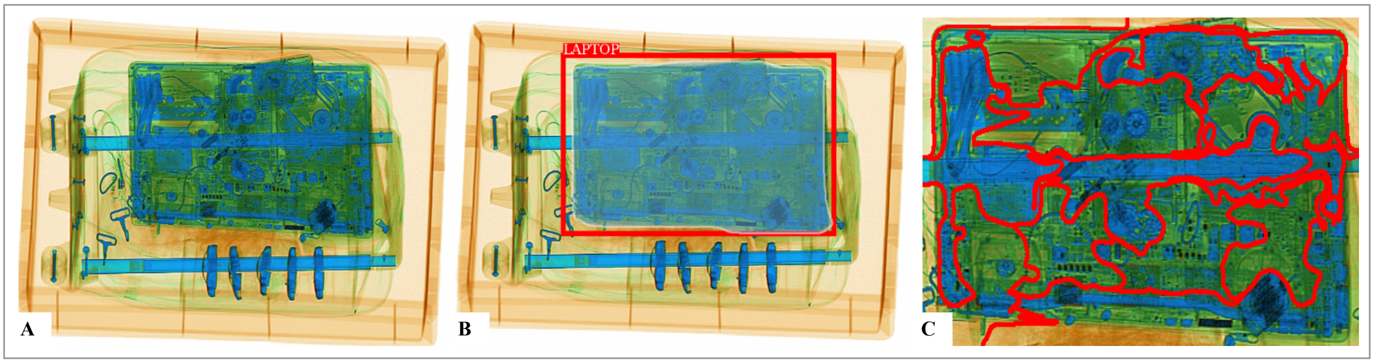

Fig. 4. Exemplar X-ray imagery (A) full frame image used for object level anomaly detection (B) via Mask R-CNN and sub-component anomaly detection (C) via SLIC over-segmentation.

components within electronic items (e.g. *'This is a laptop, but this part in the corner looks strange/unusual compared to other laptops.'* - hence refer for operator review using both full-view and as an isolated sub-component views). The potential to isolate object sub-components in this way enables efficient and effective alarm resolution at a higher granularity compared to current approaches [1]–[3].

Early work on automated threat detection wthin X-ray security images is based on hand-crafted features [4], [5], such as Bag-of-Visual-Words (BoVW), which is applied together with a classifier such as a Support Vector Machine (SVM). The original work of [4] proposes a novel variant on the BoVW model for X-ray object classification in this domain which significantly outperforms (with 99.07% true positive for Firearms detection) prior works. More recent work [6]–[8], that specifically leverage recent advances in CNN deep learning architectures [9]–[11], have now been shown to outperform these BoVW approaches in terms of true positive detection, false alarm rate (98.6% true positive / 0.2% false positive) and the range of objects that can be detected in a side by side comparison. Early work on Convolutional Neural Networks (CNN) in X-ray imaging [12] explores the use of transfer learning from another network trained on a classification task. Work considering multiple view X-ray threat detection is in its infancy using either a combination of ad-hoc 3D geometrical reconstruction [13] (lesser detection at 90-92%) or feature-space fusion pre-classification [14] (which combines features regardless of potential inter-occlusion). By contrast, the work of [7] operates at scan rate (<1 sec. per image), with higher accuracy and targets probabilistic item localization within each X-ray view. [15] has considered a unique feature representation as a critical component for detection within cluttered X-ray imagery for anomaly detection. In the works of [3], [16], unsupervised anomaly detection strategies are proposed based on high reconstruction errors produced by a generator network adversarially trained on non-anomalous (benign) stream-of-commerce X-ray imagery only.

In our proposed work, we use automatic object segmentation algorithms enabled by deep Convolutional Neural Networks (CNN, e.g. Mask R-CNN [17], Faster R-CNN [18]) together with the concept of image over-segmentation (SLIC [19]) to the sub-component level and then apply CNN classifiers ([10], [20]) to classify {*anomaly/threat, benign*} object or object sub-component of X-ray security imagery.

II. Overview of Proposed Deep Neural Network Strategies

In this work, we evaluate three different strategies for concealed threat detection in electronics item: a) full X-ray image without object localization (Fig. 4A), b) object level detection and segmentation (Fig. 4B), and c) object over-segmentation (Fig. 4C). This is followed by a binary, $\{anomaly/threat, benign\}$, classification task using deep learning based CNN classifiers.

Object detection and segmentation: We consider a number of CNN architectures, such as Mask R-CNN [17], Faster R-CNN [18], for object detection and segmentation task to explore their applicability and performance for generalised object detection within the context of X-ray security imagery. Mask R-CNN [17] relies on region proposals followed by ROI-Pooling to produce standard-sized outputs suitable for input into a secondary classifier. Mask R-CNN combines object localization with instance segmentation of the object in the image (Fig. 4B). This architecture [17] is evaluated over an electronics item (e.g. laptop) packed within cluttered X-ray security baggage imagery, for detection and segmentation task.

Object over-segmentation: For object over-segmentation (Fig. 4C), we apply Simple Linear Iterative Clustering (SLIC) [19] approach. SLIC performs iterative clustering, where initially image is segmented into roughly equal sized segments. To measure the similarity between the segments, it introduces a new distance metric which considers the size of the segment.

Classification strategy: In the final classification task, {*anomaly/threat, benign*}, applied either on full frame, object level or object over-segmented imagery, we rely on transfer learning both from a set of seminal CNN object classification networks (e.g. ResNet [10], VGG-16 [21], etc.) pre-trained on ImageNet [22]. Training is performed via transfer learning using stochastic gradient descent with a momentum of 0.9, a learning rate of 0.001, a batch size of 64 and categorical cross-entropy loss. All networks are trained on NVIDIA 1080 Ti GPU via the PyTorch framework [23].

III. Results

For evaluation, we compare the $\{anomaly/threat, benign\}$ classification (via CNN) performance of full frame X-ray image, object level segmentation and object over-segmentation. For our experimentation, our dataset (*DEEi1 - Durham Electrical and Electronics Items one-class*) is constructed using a 2D X-ray scanner with associated false colour materials mapping via dual-energy. All X-ray imagery is gathered locally by using a Gilardoni dual-energy X-ray scanner (FEP ME 640 AMX)[8]. The dataset consists of large consumer electronics items (e.g. laptop) with and without anomaly/threat (e.g. marzipan, screws, metal plates, sharps, etc.) concealment present. In total, for object over-segmentation, 7871 X-ray imagery (70 : 30 data split) and testing reported over a dataset of 4177 X-ray imagery laptop. To address the class imbalance problem, we perform data augmentation (rotation, flipping, etc.) of the anomaly/threat images. Our model performances are evaluated in terms of Accuracy (A), Precision (P), F-score (F1%), True Positive (TP%), and False Positive (FP%), as presented in the Table I.

TABLE I
Full frame vs Object level vs Object over-segmentation classification via varying CNN networks.

Object	Network configuration	A	P	F1	TP(%)	FP(%)
Full Frame	ResNet$_{50}$ [10]	**0.85**	**0.81**	**0.82**	**93.92**	**21.10**
	VGG-16 [21]	0.74	0.68	0.73	89.64	40.89
Object level segmentation	ResNet$_{50}$ [10]	**0.86**	**0.85**	**0.85**	**97.29**	**16.59**
	VGG-16 [21]	0.79	0.70	0.75	94.26	39.47
Object over-segmentation	ResNet$_{50}$ [10]	**0.97**	**0.95**	**0.97**	**98.99**	**4.54**
	VGG-16 [21]	0.94	0.92	0.93	95.89	8.55

We observe from Table I, overall object over-segmentation strategy (via Mask R-CNN [17] and SLIC [19]) outperforms full frame and object level segmentation after performing $\{anomaly/threat, benign\}$ classification. The best performing classifier for object over-segmentation (Table I lower) is ResNet$_{50}$ [10] with the highest TP and lowest FP (TP: 98.99%, FP: 4.54%). When we consider full frame X-ray images for classification task, the FP is quite high, FP: 21.10%, with the best performing ResNet$_{50}$ (Table I upper) configuration. The subtle threat/anomaly is present only in a certain regions within the object. As a result, the network performs poorly when processing the full frame images without any form of localization via a

[8] https://www.gilardoni.it/en/security/x-ray-solutions/automatic-detection-of-explosives/fep-me-640-amx/

segmentation approach. In the object-level strategy, where the target object is first detected, localized and isolated via segmentation prior to binary classification, the accuracy is improved (A: 0.86, TP: 97.29% with ResNet$_{50}$, Table I middle), due to the more focused feature representation. The over-segmented object provides higher granular information compared to the other two strategies, henceforth achieves the best performance among all strategies.

To the best knowledge of the authors the proposed work on $\{anomaly/threat, benign\}$ detection within concealed electronics items, using three different strategies, is first of its kind. As there is no prior related work is available in the literature of X-ray security imagery (e.g. object over-segmentation level classification), we are unable to compare our strategies with any existing algorithm and present the result as the benchmark.

Examples of the detection (object level segmentation via Mask R-CNN) and classification of the consumer electronics containing an anomaly are depicted in Fig. 5A. In Fig. 5B, exemplary qualitative results of over-segmentation (via SLIC) and classification of electronics item, where red colour indicates threat/anomalous region while green represents benign sub-components.

IV. Conclusion

We evaluate the performance impact of three different strategies, full frame, object segmentation, and object over-segmentation, for concealed threat/anomaly detection within consumer electronics item. Our experimental results exhibit that the best performance ($<$5% FP and $\sim$99% TP) is achieved with object over-segmentation strategy on CNN classifier.

"Given enough eyeballs, all bugs are shallow." - Eric Raymond.

Security is always a *'people, process, and technology'* solution. Here we present a deep learning enabled approach to address the challenges of threat concealment detection within consumer electronics items via a generalized approach to anomaly, rather than explicit threat, detection. In deployment such work work uniquely balance technology with human participation - *by detecting anomalous occurrences and reporting them for human review*.

Within the context of electrical items, this work offers the automatic first-stage screening of aviation baggage for anomalous electronic item detection at the component level as an indicator of potential threat presence. If an anomaly is detected in this process, then it is referred to security operators for further review. In future, this work will target more varied electronic and electrical items across a full the range of operational X-ray characteristics with future potential to span passenger (and freight) screening operations across aviation, rail, postal and maritime.

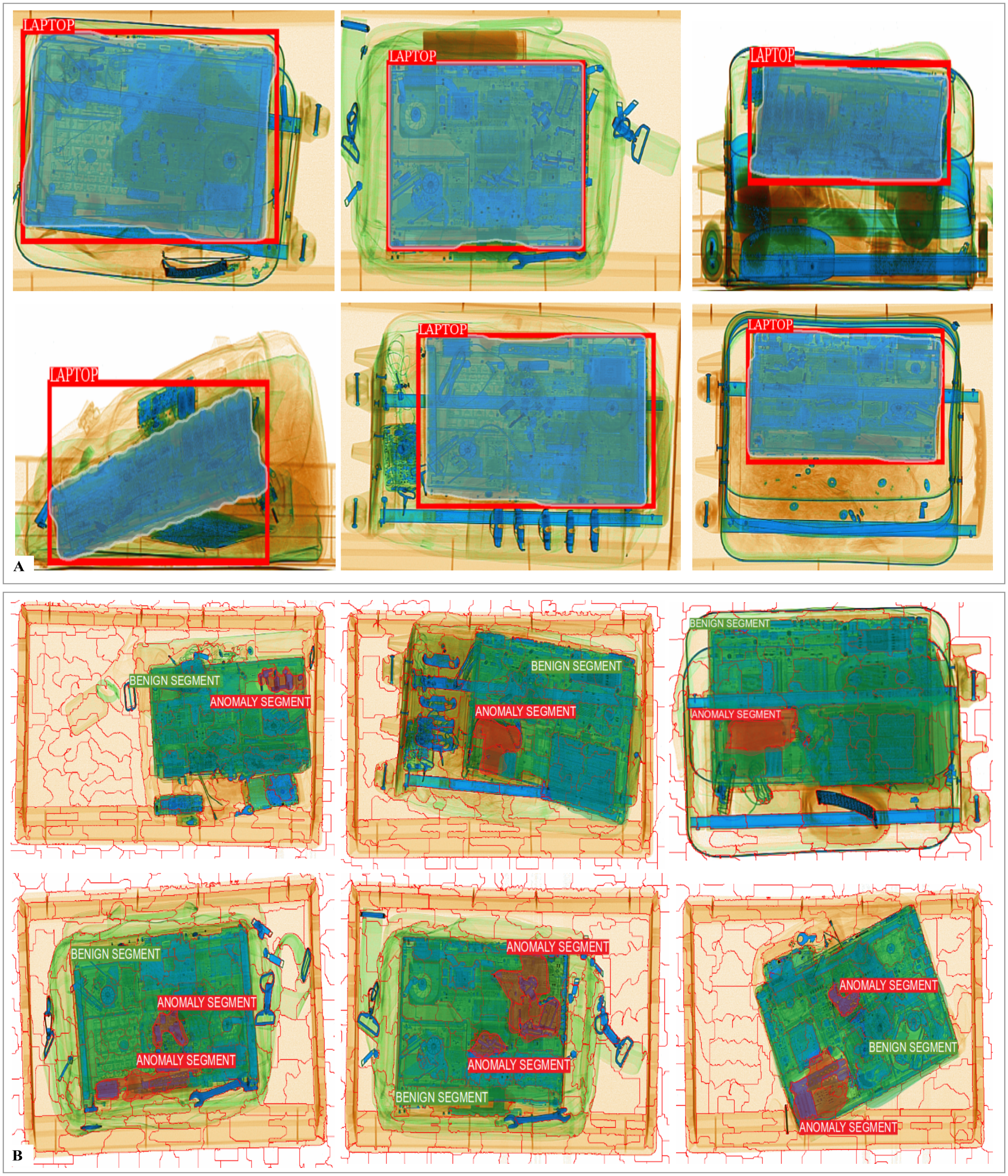

Fig. 5. Examples of $\{anomaly/threat, benign\}$ consumer electronics (laptop) detection and classification in X-ray security imagery: object level segmentation (A) by Mask R-CNN [17] and object over-segmentation (B) by SLIC [19].

References

[1] J. T. A. Andrews, E. J. Morton, and L. D. Griffin, “Detecting Anomalous Data Using Auto-Encoders,” *International Journal of Machine Learning and Computing*, vol. 6, no. 1, pp. 21–26, 2016. I

[2] L. D. Griffin, M. Caldwell, J. T. A. Andrews, and H. Bohler, “unexpected item in the bagging area,” *IEEE Transactions on Information Forensics and Security*, pp. 1–1, 2018. I

[3] S. Akcay, A. A. Abarghouei, and T. Breckon, “Ganomaly: Semi-supervised anomaly detection via adversarial training,” in *Asian Conference on Computer Vision – ACCV*. Springer International Publishing, 2018. I

[4] D. Turcsany, A. Mouton, and T. P. Breckon, “Improving feature-based object recognition for x-ray baggage security screening using primed visualwords,” in *IEEE International Conference on Industrial Technology*, Feb 2013, pp. 1140–1145. I

[5] M. Baştan, W. Byeon, and T. M. Breuel, “Object recognition in multi-view dual energy x-ray images.” in *British Machine Vision Conference*, vol. 1, no. 2, 2013, p. 11. I

[6] T. W. Rogers, N. Jaccard, E. J. Morton, and L. D. Griffin, “Automated x-ray image analysis for cargo security: Critical review and future promise,” *Journal of X-ray Science and Technology*, vol. 25, no. 1, pp. 33–56, 2017. I

[7] S. Akcay and T. P. Breckon, “An evaluation of region based object detection strategies within x-ray baggage security imagery,” in *2017 IEEE International Conference on Image Processing*, Sep. 2017, pp. 1337–1341. I

[8] 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 Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2203–2215, Sep. 2018. I

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017. I

[10] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2016, pp. 770–778. I, II, I, III

[11] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” in *AAAI Conference on Artificial Intelligence*, 2017. I

[12] 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 *IEEE International Conference on Image Processing*, Sept 2016, pp. 1057–1061. I

[13] D. Mery, V. Riffo, I. Zuccar, and C. Pieringer, “Automated x-ray object recognition using an efficient search algorithm in multiple views,” in *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, June 2013, pp. 368–374. I

[14] M. Baştan, “Multi-view object detection in dual-energy x-ray images,” *Machine Vision and Applications*, vol. 26, no. 7, pp. 1045–1060, Nov 2015. [Online]. Available: https://doi.org/10.1007/s00138-015-0706-x I

[15] L. D. Griffin, M. Caldwell, J. T. Andrews, and H. Bohler, “unexpected item in the bagging area: Anomaly detection in x-ray security images,” *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 6, pp. 1539–1553, 2019. I

[16] S. Akcay, A. A. Abarghouei, and T. Breckon, “Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection,” in *2019 International Joint Conference on Neural Networks*, 2019. I

[17] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *Proc. of the IEEE Int. Conf. on Computer Vision*, 2017, pp. 2961–2969. I, II, III, 5

[18] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” in *Advances in Neural Information Processing Systems*, 2015, pp. 91–99. I, II

[19] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Ssstrunk, “Slic superpixels compared to state-of-the-art superpixel methods,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, Nov 2012. I, II, III, 5

[20] Y. Wang, V. I. Morariu, and L. S. Davis, “Learning a discriminative filter bank within a cnn for fine-grained recognition,” *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4148–4157, 2018. I

[21] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *International Conference on Learning Representations*, 2015. II, I

[22] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, “Imagenet large scale visual recognition challenge,” *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, Dec 2015. II

[23] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, “Automatic differentiation in pytorch,” 2017. II