International Journal of Engineering and Management Research e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume- 9, Issue- 3 (June 2019)
www.ijemr.net https://doi.org/10.31033/ijemr.9.3.03

Passenger Screening using Deep Learning and Artificial Neural Networks

Sarthak Arora[1], Ankit Singh Rathore[2] and Saurabh Gautam[3]
[1]Student, Department of Information Technology, Maharaja Agrasen Institute of Technology, GGSIPU, INDIA
[2]Student, Department of Information Technology, Maharaja Agrasen Institute of Technology, GGSIPU, INDIA
[3]Assistant Professor, Department of Information Technology, Maharaja Agrasen Institute of Technology, GGSIPU, INDIA

[1]Correspondence Author: sarthakarora.1997@gmail.com

ABSTRACT

In this research, we have to detect the contrabands hidden in the human body's scanned images at airport security machines using segmentation and classification. Present algorithm of security scanning machines at the airports of USA are producing high rate of false negatives which in cases lead to engage in a secondary, manual screening process that slows everything down. So to resolve this problem and to improve the detection of contrabands, new and efficient algorithm need to be made.

***Keywords*--** Transportation Security Administration, Advanced Image Technology, CNN

I. INTRODUCTION

Use of Neural networks and deep learning in airport security is a major breakthrough. Deep learning can provide new capabilities and approaches for addressing security problems. Airport security refers to the techniques and methods used in an attempt to protect passengers, staff and planes which use the airports from accidental/malicious harm, crime, and other threats. Aviation security is a combination of human and material resources to safeguard civil aviation against unlawful interference. Unlawful interference could be acts of terrorism, sabotage, threat to life and property, communication of false threat, bombing, etc. Large numbers of people pass through airports every day. This presents potential targets for terrorism and other forms of crime because of the number of people located in one place. Similarly, the high concentration of people on large airliners increases the potentially high death rate with attacks on aircraft, and the ability to use a hijacked airplane as a lethal weapon may provide an alluring target for terrorism. So to resolve this problem and to improve the detection of contrabands, new and efficient algorithm need to be made. For this, dataset has been downloaded from kagglehttps://www.kaggle.com/c/passenger-screening-algorithm-challenge/data, which was provided by DHS (Department of Homeland Security) to kaggle. The dataset contains large number of body scans obtained by new generation of millimetre wave scanner called High Definition-Advanced Imaging Technology (HD-AIT) system. The scans comprise of subjects who are wearing different clothing types, having different BMI's, different number of contrabands or threats and different types of contrabands. To pre-process the images and to train our CNN model, python is used as a prominent programming language. During pre-processing of dataset's images , cropping and segmentation of images of about 1200 bodies has been done and labelled those segmented and cropped regions as a train data and test data. The preprocessed data are pipelined to CNN deep learning model to predict probability of contrabands in predefined 17 body zones. According to Bart Elias, Transportation Security Administration (TSA) has deployed 1,800 units Advanced Image Technology (AIT) throughout the airports on US and the acquisition cost per unit is about $175,000. To fully implement this system, the overall annual cost for purchasing, installing, staffing, operating, supporting, upgrading and maintaining sum to about 1.17 billion. In addition, the Department of Homeland Security's (DHS) in USA has found a higher false alarm that is produced from the algorithms which is used in today's scanners at the airports. Although, the investment being made represents a huge amount of money, it is not solving the security problem with regarding to the detection of threats. The Security still needs to develop technology to fix these algorithms to minimize the errors. Like others cleaners of data, working with images need a special care in filter data that should be a ground truth as possible. Many difficulties are aroused in filtering the noise in the images and in segmenting the images in regions that contains the object. Errors in the segmentation can mislead the security system and generally produces false results. The ideal is to reduce the noise and produce some threshold that correct the segmentation regions. Considering this scenario, this project will present a model to do the classification of the regions of the body´s images using supervised machine learning to identify the hidden threats hired in the human bodies. The results will be analyzed to verify the true or false of the information collected in the model. The objectives proposed in this project are: Produce an algorithm that fractionated the human body image in regions to be able to identify the body's region correctly.

II. LITERATURE REVIEW

While long lines and frantically shuffling luggage into plastic bins isn't a fun experience, airport security is a critical and necessary requirement for safe travel.

No one understands the need for both through security screenings and short wait times more than U.S. Transportation Security Administration (TSA). They're responsible for all U.S. airport security, screening more than two million passengers daily.

Whenever TSA's sensors and algorithms predict a potential threat, TSA staff needs to engage in a secondary, manual screening process that slows everything down. And as the number of travelers increase every year and new threats develop, their prediction algorithms need to continually improve to meet the increased demand.

Currently, TSA purchases updated algorithms exclusively from the manufacturers of the scanning equipment used. These algorithms are proprietary, expensive, and often released in long cycles. In this competition, TSA is stepping outside their established procurement process and is challenging the broader data science community to help improve the accuracy of their threat prediction algorithms. Using a dataset of images collected on the latest generation of scanners, participants are challenged to identify the presence of simulated threats under a variety of object types, clothing types, and body types. Even a modest decrease in false alarms will help TSA significantly improve the passenger experience while maintaining high levels of security.

III. METHODOLOGY

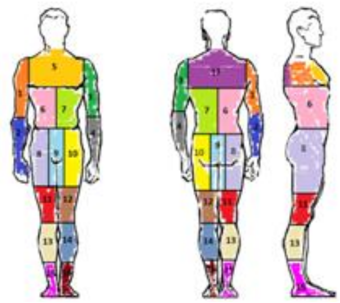
DataSet Available

The dataset used in this project comes from DHS and Kaggle competition at: https://www.kaggle.com/c/passenger-screening-algorithm-challenge/data. This dataset contains a large number of body scans acquired by a new generation of millimeter wave scanner called the High Definition-Advanced Imaging Technology (HD-AIT) system. The competition task is to predict the probability that a given body zone (out of 17 total body zones) has a threat present. The images in the dataset are designed to capture real scanning conditions. They are comprised of volunteers wearing different clothing types (from light summer clothes to heavy winter clothes), different body mass indices, different genders, different numbers of threats, and different types of threats. Due to restrictions on revealing the types of threats for which the TSA screens, the threats in the competition images are "inert" objects with varying material properties. These materials were carefully chosen to simulate real threats

Approach

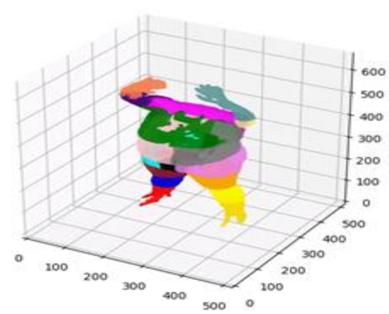
1. Analyzing Threat Zones

The scans to be analyzed in this project all segment the body into numbered "threat zones". Our modeling and results will have to give probability of contraband within a given threat zone. It included a visualization of the threat zones and the corresponding numbering scheme.



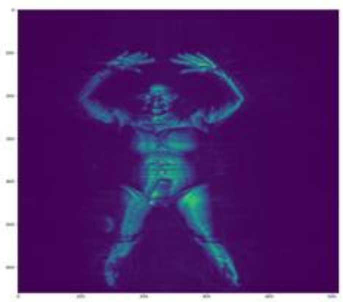
2. Segmenting and Cropping the Scans

Scans are images with views from shots taken at a regular interval of 22.5-degree rotation, 360 degrees around the subject scanned. So for each subject, we'll get 16 scans. Each scan will contain a view from a given angle of any visible threat zones. Part of our approach in this Project is to isolate each individual threat zone from every visible angle. Later, we want to be able to make features out of each individual threat zone from each angle that a given threat zone is visible. That will allow us to later train on each threat zone individually from every view in a 2D format. Each image is divided into "sectors", the sector numbers used below correspond to the threat zone number. But as the images in the sequence rotate around the subject, threat zones may be visible from another sector. Below, we'll use these vertices to capture isolated views of a give threat zone



3. Threat Zone Slice Lists and Crop Lists

For each zone, we pick the sector with the best view of that threat zone. Since there are 16 images per subject, each zone's slice list contains 16 entries each naming the sector where the given threat zone is best visible. At some angles, a given threat zone may not be visible. We just set those values to None so that it's easy to iterate the list and ignore the scans where a given zone is not visible. We call this a threat zone slice list. To create a feature for a given threat zone. We'll iterate this list and use the vertices to establish a region of interest and mask off the threat zone. We used the same technique to then crop the segmented images to reduce the size. Each element in the zone slice list contains the sector to use in the call to roi().



4. The Preprocessor

Passengers who are being scanned for contraband are referred to as "subjects". Our approach in this project is to isolate each individual threat zone from every visible angle and then make features out of each individual threat zone from each angle that a given threat zone is visible. This allows us to train on each threat zone individually from every view in a 2D format. The preprocessor loops through the data one subject at a time, transforms the images, isolates threat zones, and uses a set of vertices to crop each image to 250x250. Images are saved in mini batches by threat zone, so that they can be read into the trainer.

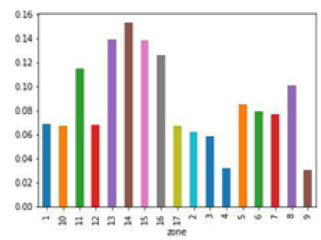
IV. RESULT

In this Passenger Screening Algorithm, we have trained our model with about preprocessed images of having threat probability and preprocessed images of having non-threat probability. The accuracy obtained during training is 95.04% and the loss is about 0.0414. After the training phase, Passenger Screening Algorithm is tested over subject images to find the probability and threat percentage in each body zone.

Probability of threat in each body zone

V. CONCLUSION

To address the ever-evolving threats to airport security, we present the most prominent and enhanced model to detect the concealed threats in the human body by passenger screening algorithm. In this project, it was developed a method to do the segmentation of the body in zones. It was presented 17 zones to do the learning process. Using robust module CNN and ResNet, the model could learn the zones with threat or not. Combining all these, the model could classify those zones of the bodies with a robust result that can help One important point to be considered here is the fact that the model worked for the trained images of the sample used, it was not considered any external image. The results produced by the algorithm is highly accurate for predict zones, it can indicate an over fitting of learning. Thus outperforming traditional methods of threat detection.



Bar graph showing mean probability of having threat in each zone.

REFERENCES

[1] M. Barreno, B. Nelson, A. D. Joseph, & J. D. Tygar. (2010). The security of machine learning. *Machine Learning*, *81*, 121–148. Available at: https://people.eecs.berkeley.edu/~tygar/papers/SML/sec_mach_learn_journal.pdf.

[2] B. Nelson, B. I. P. Rubinstein, L. Huang, A. D. Joseph, S. J. Lee, S. Rao, & J. D. Tygar. (2012). Query strategies for evading convex-inducing classifiers. *Journal of Machine Learning Research, 13*, 1293–1332.

[3] L. Bottou. (2012). Stochastic gradient descent tricks. *In Neural networks: Tricks of the trade. Springer*, 421–436.

[4] S. Sarraf & G. Toghi. (2016). Deep learning-based pipeline to recognize alzheimer's disease using fmri data. *In Future Technologies Conference (FTC). IEEE*, 816–820.

[5] S. V. D. Walt, S. C. Colbert, & G. Varoquaux. (2011). The numpy array: a structure for efficient numerical computation. *Computing in Science & Engineering, 13*(2), 22–30.