Journal of Knowledge Learning and Science Technology
ISSN: 2959-6386 (Online)
2024, Vol. 3, No. 4, pp. 372–382
DOI: https://doi.org/10.60087/jklst.v3.n4.p372

Research Article

Concealed Object Detection and Localization in Millimetre Wave Passengers' Scans

Maher Gerges[1], Ahmed Elgalb[2], Abdelrahman Freek[3]

[123]Independent Researcher, USA

Abstract

The exponential growth in air travel has heightened airport security risks, making traditional manual screening methods increasingly inefficient. To address this challenge, we propose an automated system for detecting and localizing concealed prohibited objects in millimeter-wave images of passengers. Our approach leverages image processing and data mining techniques to enhance detection accuracy and efficiency.

In our method, each millimeter-wave image undergoes preprocessing to filter out noise and artifacts. The images are then segmented into zones, each treated as an individual image. Zero-centering and normalization are applied to these zones to optimize the performance of the neural network. The dataset is divided into training and testing sets, with the training set shuffled to improve learning. The neural network is trained on this data to predict the presence of potential threats.

This paper presents the methodology of our solution and discusses how it addresses the challenges of passenger scanning and false alarms. Preliminary results indicate that our system has the potential to significantly improve detection rates while re-ducing unnecessary alerts, thereby enhancing overall airport security efficiency.

Keywords

Airport security, Automated system, Concealed objects, Millimeter-wave images, Image processing, Neural network, Detection accuracy, False alarms, Security efficiency

1. Introduction

The rapid increase in air travel has heightened security risks at airports, particularly concerning the concealment of prohibited objects under passengers' clothing. Inefficient passenger screen-ing methods have contributed to security breaches, including bombings over the years. With the advancements in technology, there is a pressing need to improve passenger screening processes by developing automated target recognition systems that are both time-efficient and accurate.

Passengers and airport staff often experience delays and false alarms due to current screening methods, which can be both inconvenient and disruptive. Therefore, the demand for a more effective and precise passenger screening system has

*Corresponding author: Maher Gerges

Email addresses:

maher2018fayez@gmail.com

Received: 01-10-2024; **Accepted:** 01-11-2024; **Published:** 25-12-2025

become essential. Various technologies have been adopted to address this issue, such as X-ray backscatter imaging and millimeter-wave imaging. Figure 1 illustrates a full-body scan using X-ray backscatter technology [2].

However, the use of X-rays in passenger screening raises concerns regarding privacy and potential health risks due to ionizing radiation. As a result, millimeter-wave imaging has emerged as a safer and more acceptable alternative. Applying image processing and data mining techniques to millimeter-wave images is currently one of the most effective methods for detecting concealed prohibited items under passengers' clothing. Figure 2 shows an example of millimeter-wave imaging technology in use. [1]

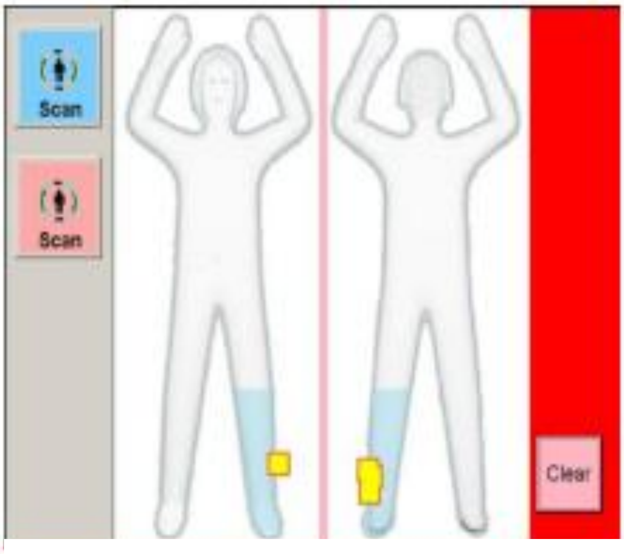


Figure 1. *The use of Automated Target Recognition in airports to overcome the problem of privacy. The figure shows how only and outline of the body is displayed and the location of the threats is highlighted, without showing any details of the body [1]*

Due to these issues, millimeter-wave technology has gained prominence in passenger screening, while X-ray systems continue to be used primarily for luggage inspection. Millimeter-wave imaging is effective in detecting concealed objects because of its reflective characteristics and ability to penetrate fabrics without the harmful effects associated with X-rays. However, detecting objects of various shapes and poses remains challenging.

To mitigate privacy concerns, algorithms like Automated Target Recognition (ATR) have been implemented in millimeter-wave imaging systems. ATR replaces detailed body images with generic outlines, highlighting only the positions of potential hidden threats, as illustrated in Figure 3 [1]. This approach enhances privacy while maintaining security effectiveness.

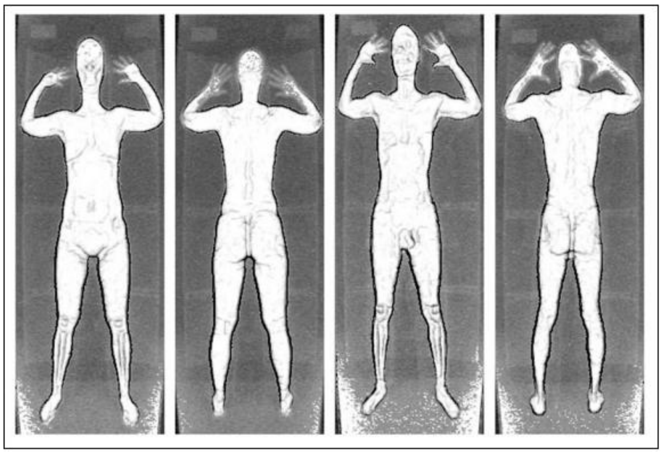


Figure 1 *Full body scan by X-Ray Backscatter after applying some privacy algorithms. The left image is an image for a female while the right one is an image for a male. [1]*

Despite advancements in imaging technologies and the incorporation of image pro-cessing techniques, passenger screening processes remain inefficient in terms of time and accuracy. Travelers often face long wait times, and false alarm rates are still high. Therefore, there is a pressing need to improve these systems by employing ad-vanced image processing and data mining techniques to enhance detection accuracy and reduce screening times.

2. Scope and Objectives

This paper presents an algorithm for detecting prohibited items concealed under passengers' clothing using millimeter-wave images and image processing techniques. The algorithm is designed to predict and localize threats from concealed objects, uti-lizing a dataset provided by Kaggle. The focus of this paper is on the image pro-cessing methods that enhance object detection in millimeter-wave images of the hu-man body, without delving into hardware aspects.

3. Paper Organization (Structure)

The paper is organized as follows:

Chapter 1: Introduction and Related Work

Introduces the topic and reviews four different methods previously published on de-tecting and localizing objects in millimeter-wave images. This chapter includes a graph illustrating the various processes each image undergoes in these methods until the final output.

Chapter 2: Methods and Materials

Details the dataset used in the proposed solution and describes the different processes that images undergo in the solution.

Chapter 3: Results, Experiments, and Accuracy

Presents the experimental results and evaluates the accuracy of the neural network in detecting and localizing concealed threats.

Chapter 4: Conclusion

Summarizes the findings and discusses the implications of the work.

3.1 Work Methodology

Advancements in image processing and data mining techniques have revolutionized object detection algorithms. In this solution, we utilize a dataset of millimeter-wave images captured using a new technology called High Definition-Advanced Imaging Technology (HD-AIT). The images depict real people with concealed objects of various shapes and types under their clothing.

The methodology involves several key steps:

Preprocessing: All images undergo preprocessing to remove noise and artifacts, enhancing image quality and preparing the data for further analysis.

Zone Segmentation: After preprocessing, each image is cropped into smaller zones and saved as separate images. This segmentation allows for more precise analysis, as each zone is treated individually.

Labelling: Each zone is labelled to indicate whether it contains a threat or not. This labelling is crucial for supervised learning in the neural network.

Data Splitting: The labelled zones are divided into training and testing sets. The training set is used to train the neural network, and the testing set is used to evaluate its performance.

Neural Network Training: A neural network is trained using the training set to detect and localize concealed threats within the zones.

Testing and Evaluation: The testing set is used to assess the accuracy and effective-ness of the neural network, measuring its ability to correctly identify threats.

By focusing on these image processing techniques and leveraging the HD-AIT da-taset, the proposed algorithm aims to improve the detection and localization of pro-hibited items in millimeter-wave images of the human body.

3.2 Work Plan (Gantt chart)

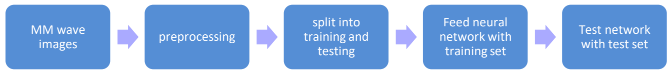


Figure 4 *shows the workflow of the proposed solution. Only the main processes are shown. The internal expla-nation and sub-processes will be shown later in the materials and methods section.*

3.3 Related Work (State-of-the-Art)

In a study presented at the 2016 7th International Conference on Cloud Computing and Big Data [2], the authors proposed an object detection method comprising four main steps: pre-processing, image division, feature extraction and re-encoding, and classifier training. Millimeter-wave images often contain noise and artifacts due to factors such as background scattering, which can lead to decreased system accuracy. Additionally, processing three-dimensional (3D) images requires more computational time compared to two-dimensional (2D) images. Therefore, transforming images from 3D to 2D during pre-processing is essential to reduce computational time and remove noise. Standard 2D image filters effectively decrease circuit noise, while system calibration can eliminate artifacts. Since different parts of the human body vary in shape and pose, dividing the body image into segments is necessary. For example, the brightness and contrast in the torso area are higher than those in the legs and arms.

The grayscale intensity of hidden objects differs from that of the human body in millimeter-wave images due to reflectivity disparities. However, because hidden objects may have various shapes and poses, it is crucial to identify image descriptors that are invariant to transformations. Considering these factors, the method emphasizes the importance of saliency detection, Histogram of Oriented Gradients (HOG), and Scale-Invariant Feature Transform (SIFT) features. Both HOG and SIFT represent the gradient structures that form shapes within the image. The Itti and Spectral Residual models are prominent saliency detection models, with the Spectral Residual model being independent of object characteristics and features. This model utilizes the logarithmic spectrum of an image to extract the spectral residual in the spectral domain and constructs the corresponding saliency map in the spatial domain. In real-world scenarios, hidden objects have diverse materials, forms, and poses, while the image backgrounds typically exhibit similar brightness, making the Spectral Residual model a better choice for saliency detection.

Traditionally, HOG and SIFT features are directly fed into a Support Vector Machine (SVM) classifier for classification. According to [2], this approach is inefficient for millimeter-wave image classification due to the complexity of the data. To achieve a more efficient method, a two-layered model using sparse coding is employed to characterize the image features. Multi-layer models can represent more detailed aspects of millimeter-wave images. By re-encoding features through sparse coding, reconstruction errors are minimized and computational complexity is reduced.

Considering computation time, the Synthetic Targets Detection Method utilizes a linear SVM as the classifier. Experiments comparing linear and non-linear SVM classifiers indicate that although non-linear SVMs provide more accurate recognition, they also result in higher complexity and longer processing times. To enhance recognition speed, a linear SVM is selected as it offers a better balance between accuracy and efficiency. Using linear SVM classification, the method achieves a total processing time of approximately 600 milliseconds per image, making it feasible for real-time detection systems.

In another study [3], the authors follow similar steps to the first method but implement different filtration and classification techniques, and select different features. For pre-processing, both linear and non-linear smoothing filters are applied. A statistical filter replaces each pixel value with the average of a random sample of neighboring pixel values to reduce noise. Subsequently, a 5×5 median filter is used to remove boundary artifacts by traversing the image with a 5×5 matrix, taking the median value of the matrix, and replacing the current pixel value.

Haar-like features and Local Binary Patterns (LBP) are chosen as image feature descriptors. For classification, the authors argue that applying a single classifier is inefficient due to the varying sizes of classes. Instead, a group of six binary classifiers is applied, including logistic regression with variable transformations, Support Vector Machine, Random Forest, Extremely Randomized Trees, and AdaBoost. The results demonstrate an accuracy of 68% for both classes, indicating that the ensemble of classifiers improves the classification performance.

In [4], the authors implement multi-level segmentation using K-means clustering along with the Expectation-Maximization (EM) algorithm and the Bayesian decision rule. Edge detection using the Canny method is employed to represent image features. Two levels of segmentation are applied: the first is global segmentation, which separates the body from the background, and the second is local segmentation, which isolates the concealed objects from the body. According to the paper, applying multi-level segmentation increases the detection accuracy by effectively distinguishing between the body, background, and concealed objects.

Another study [5] applies Iterative Steering Kernel Regression (ISKR) and Local Binary Fitting (LBF) for denoising and segmentation to separate objects from the body. Additionally, Non-Local Means (NL-means) is used to denoise the images. The results show that the correctness of full-body segmentation using NL-means and LBF was 83%, which increased to 86% when applying ISKR with LBF. For threat segmentation, the correctness improved from 79% with NL-means and LBF to 90% with ISKR and LBF. The precision of threat detection was 88% using NL-means and LBF, increased to 93% with ISKR and LBF, and reached 100% when applying NL-means and the EM algorithm. These findings suggest that advanced denoising and segmentation techniques significantly enhance the accuracy and precision of concealed object detection in millimeter-wave images.

3.4 Analysis of the Related Work

All the methods presented rely on image processing techniques. Each of the four methods preprocesses images to remove noise and artifacts. Three of these methods use feature extraction to represent objects, followed by a linear classifier to predict the presence of hidden items based on the extracted features. The fourth method em-ploys multi-level segmentation. However, none of these methods has utilized a full dataset to test and demonstrate the exact accuracy of their approaches, primarily be-cause the relevant dataset was only published by Kaggle in 2017. Consequently, these methods did not incorporate convolutional neural networks (CNNs), which require large datasets to achieve high prediction accuracy.

The features extracted in the presented methods may effectively describe the shape and size of hidden objects, but they are insufficient for predicting new concealed items. This limitation is why we propose using a convolutional neural network in our solution. Traditional feature extraction methods perform well when objects are clear and their features are well-known, but their effectiveness diminishes when features are ambiguous. Similarly, multi-level segmentation relies heavily on edge detection and works best when object edges are clear and easily detectable. In the context of millimeter-wave images, where objects may not have well-defined edges, this meth-od could result in a high rate of false alarms.

Conversely, convolutional neural networks can extract hundreds of features in their feature extraction layers, providing a more detailed and accurate description of ob-jects. This capability enhances the accuracy of object detection in millimeter-wave images, making CNNs a more suitable choice for detecting concealed prohibited items under passengers' clothing.

3.5 Solution Methodology

There are different techniques that have been adapted to solve the problem of detecting concealed objects under passengers' clothes in aviation security at airports. Four of these methods have been presented in the related work section, and their drawbacks have been discussed in the analysis. Our solution relies on image processing techniques to filter millimeter-wave images and divide them into zones. The prediction process is performed using a convolutional neural network. Each millimeter-wave image is filtered to remove noise and artifacts, and then divided into zones according to the body zones provided by Kaggle in the dataset. Cropping the entire dataset produces many zones, which are split into training and testing sets. The training set is used to train the network, and the testing set to evaluate its accuracy.

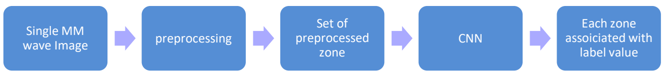


Figure 2. *This figure shows the process of evaluating single MM wave image. Starting from preprocessing the image and division into zones. Then Make the CNN check whether zone contains threat or not.*

4 The proposed method

4.1 Dataset

Passenger screening problem dataset is provided by Kaggle as a dataset for the challenge to solve the passenger screening in the airports problem. The dataset is 64 gi-gabyte of millimeter images in different formats, csv files, and body zones single pang image.

4.1.1 The millimeter wave images

These images are taken by a new millimeter wave image technology called High Def-inition-Advanced Imaging Technology (HD-AIT). These images are of real volun-teers who are different in gender, mass, width, and height wearing different kinds of cloths. Underneath the volunteers' cloths placed different threads in different loca-tions of the body. These threads are chosen to simulate real life threats that might be hidden under passengers' cloth.

4.1.2 Size of dataset

The original size of the dataset images is more than three terabytes which is stored on google cloud of four formats files (. ahi, aps, .a3d, and .a3daps). The four types of files for the same image but they are different formats of the 3D scan. Because it is hard to download and save this huge size of data, Kaggle provided a complete dataset of images but only two formats of files aps and a3daps which is smaller in size. Each file is composite of a header size 512 bytes followed by the data of the file.

4.1.3 Dataset files

.ahi files: calibrated object data files which are the biggest files format in size (2.26GB for each file).

. A3d files: 3D projection of the images. This format can be generated using the combined image algorithm that combines eight 3D scans into single 3D projected image. The size of each file is 330MB for each file.

.Aps files: projected of angle sequence of 2D images. Aps file is a sequence of 16 2D millimeter wave scans. Each scan is captured at regular angle 22.5 degree with 360 degree around the body. So the aps image is a 16 slice of the mm wave image that combined together producing kind of projected 3D image. It is like taking the front view of the body then after 22.5 degree capture the second view and so on till the completion of the 360 degrees. This format of files the smallest one in size only 10.3MB for each file. Since this type of files is the smallest and represents all the details that is need for this problem, so we will use it in our solution. Figure 5 shows the 16 2D views of the aps file.

.A3daps files: combined of sequence angle images and its size is 41.2MB for each file. It is very close to the aps files, but it is combined 3D image. It is used for better visualization though creating sequence of the a3d files with angle intervals.

Body zones .Png file: is used to describe how the image can be divided into zones. Each image is divided into 17 threat zones. As mentioned before that all the images are mm wave represented as 16 2D views images, so that each zone will have 16 2D views. The abs files for instance is a 16 view of the 2D image from different angles, then each zone of the aps file has 16 view to. Some view a certain zone is very clear others are not. Figure 6 shows the division of body into 17 zones.

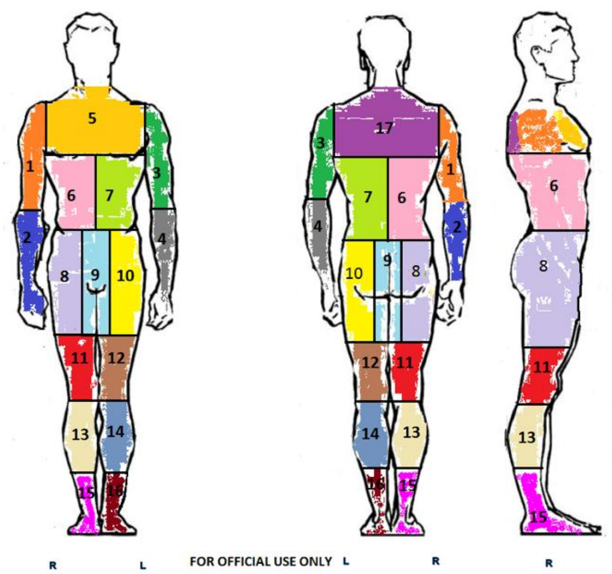


Figure 8. *This figure shows how the whole image is divided into zones. Notice that each zone is represented in form of 3D, not just o2D view. For example, zone number 6 can be viewed in the front view, back view and right view but it is hidden in the left view. [6]*

.csv file: Each image is named by a unique identifier image Id. Csv file contains for each image Id followed by zone number a label (Boolean value) indicates whether this zone contains a threat or not. For example, image 00360f79fd6e02781457eda48f85da90_Zone1 this is image with Id 00360f79fd6e02781457eda48f85da90 and zone number 1. Since the dataset contains 1,247 images and each image is divided into 17 zones, then it is supposed that the csv file contains 21,199 labels. But it only contains 19,500 labels, which means that there are some missing images.

4.2 Zone threats analysis

Figure 7 shows the number of threats in each zone individually in all images (for each zone in all subjects how many zones contains threats over total number of subjects). The chart shows that zone number 1 get the highest percentage. Which looks strange, since zone 1 represents the right upper arm and the part of the body is not the best place to hide objects. Zone number 2 comes with the second highest percentage. Which also strange because zone 2 represents the right fore arm which is not an appropriate place to hide something. The highest zone supposed to be 6 or 7 which represents right and left torso. Because this part of the body is the best for passengers to hide a certain objects or they might forget to take of their belt.

The following table show the number of threats in each zone and the exact per-centage of threats in each zone.

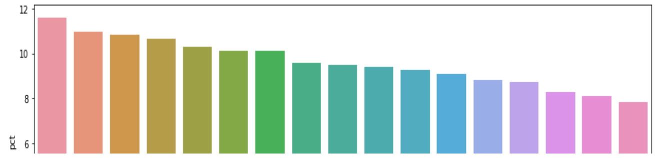


Figure 3. *This figure show the percentage of threats in each zone. Notice that zone number 1 is the highest percentage. However it represents the upper of the right arm. The highest zone supposed to be zone 6 or 7*

Zone	Number of threats	Percentage
Zone1	133	11.595%
Zone2	126	10.985%
Zone8	124	10.811%
Zone14	122	10.636%
Zone15	118	10.288%
Zone11	116	10.113%
Zone6	116	10.113%
Zone13	110	9.590%
Zone16	109	9.503%
Zone4	108	9.416%
Zone5	106	9.241%
Zone3	104	9.067%
Zone12	101	8.806%
Zone10	100	8.718%
Zone17	95	8.282%
Zone7	93	8.108%
Zone9	90	7.847%

Figure 10. *This table show each zone number, how many threats in this zone, and the percent-age of threats in this zone.*

4.2.1 Reading and plotting images

Reading images and plotting images was the first and most difficult process. The im-ages formats are not well known and handling these formats is not straight forward process. There is no built-in function in python, MATLAB, or any other pro-grammers to read these formats of images. The only way is to read the header and data of the image. The image header is 512 bytes than contains a description of the image data. Start-ing by creating a dictionary to hold the information read from the header using numby package functions. The dictionary will hold all the descriptors of image in the header like the file name, file type, dimensions, and so on.

Reading images and plotting images was the first and most difficult process. The im-ages formats are not well known and handling these formats is not straight forward process. There is no built-in function in python, MATLAB, or any other pro-grammers to read these formats of images. The only way is to read the header and data of the image. The image header is 512 bytes than contains a description of the image data. Start-ing by creating a dictionary to hold the information read from the header using numby package functions. The dictionary will hold all the descriptors of image in the header like the file name, file type, dimensions, and so on.

Next step is reading the image data using the information that stored in the header dictionary. Starting by getting the im-age dimensions form the header (X, Y, Z). Then skip 512 bytes of the header and move direct to the data using seek function. Once the data of the image has been read, it is time to plot the image. Plotting the image also cannot be done using a built-in function in python. The only way is to build a plotting function using the built-in functions from matplotlib package. As we men-tioned before, the aps files are kind of 16 2D angles scans combined creating a repre-sentation of 3D image with less size. So, to plot the image we looped over the 16 2D.

4.2.2 Retrieving single 2D view of aps file

In this step, a single angle view of the 16 views is extracted from the aps files to han-dle it as normal 2D image. This ena-bles us to apply all 2D image processing tech-niques like fil-tration, cropping, and other techniques on the single image. Then later on the same processes are applied to all 16 views.

4.2.3 Convert to grayscale

Once the image is read, each 2D view of the image is con-verted to grayscale rep-resentation. This step is performed for many reasons: first we do not care about the RGB image. In other words, there are no important features based on the col-oured image needed in this application. Second reason is com-plexity and speed, processing grayscale image is faster and simpler than handling RGB one. Converting to gray-scale transform pixel from 3D to 1D which facilitate many pro-cesses like edge detec-tion. Figure 4 shows the grayscale im-age and its histogram. Whether working on a colour image or convert to grayscale one depends on the application working on. It depends on whether we need the colour details or not. In object detection applications like passenger screening, we do not care about the RGB colours details. Because we need only the grayscale image to extract some features like edge detec-tion, SIFT, HUG, and other features that mainly depends on the grey level of the image not the colour one. So, converting image to grayscale is a major process in all object detec-tion algorithm. Figure 10 shows the front view of the image after converting to gray-scale.

4.2.4 Image enhancement

In order to get rid of the noise and artefacts caused by the reflection of human body to the millimetre waves and enhance the contrast of the image, different techniques has be adapted. The first method to remove the noise and artefacts is the threshold method with threshold minimum 12 and the new value 0 is applied to the image. The threshold results in replac-ing all pixel values less than 12 to 0 which remove almost all the artefacts and noise around the body. The threshold tech-niques is simply takes an input thremax, thremin, and newVall inputs. Then move over the image matrix and replace all the pixel values above the thremax and all the pixel values below the thremin, and replace these values with the newVal. In case there is only thremin and newVal like our problem, then all the values below thremin will be replaced by the newVal. This method works well when the noise we want to remove are less than a certain value or higher than the normal values of our images. Next example shows the result of applying threshold technique with thremin is10, thremax is 50, and the newVal is 0.

10	0	11	3	7
3	2	2	10	15
35	57	65	32	75
68	12	35	9	0
21	5	68	48	32

Apply threshold with thremax 50, thremin 10, and newVal 0

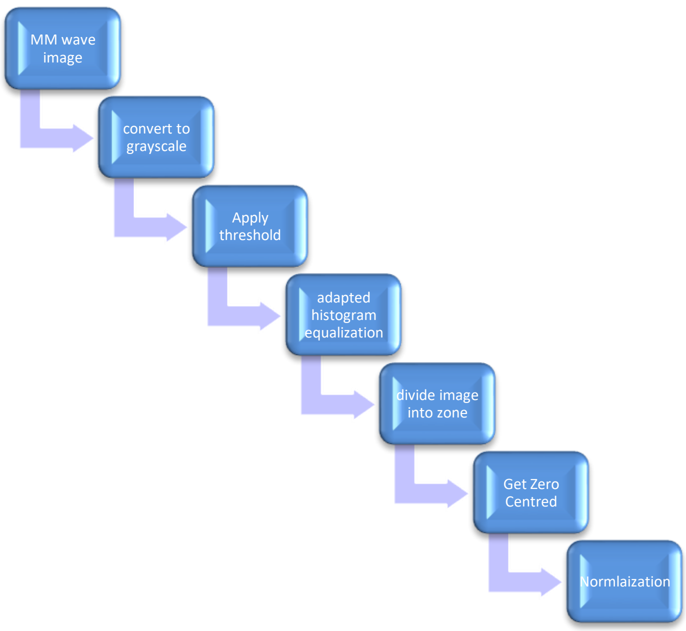
10	0	11	0	0
0	0	0	10	15
35	0	0	32	0
0	12	35	0	0
21	0	0	48	32

And then to enhance the contrast of the body, a special histogram equalization method is adapted by creating what is called clahe. This method divide image into tiles – small blocks of the image- and then apply the normal histogram equalization to each block. Also, a contrast clip limit is used so that values greater than limit are divided and distributed over the other values. Figure 9 shows the new image after applying this method of threshold.

4.2.5 Dividing image into zones

Now after converting the image to grayscale and improve its quality using special histogram equalization and threshold. Since we have for each separated zone a label value indicates whether this zone contains a threat or not. So, in this stage the image will be divided into 17 thread zones based on the body zones image provided by Kaggel in the dataset. There are many segmentation techniques that have been pro-posed for 2D image segmentation. But the 3D segmentation for signal images is so-phisticated process and it takes a long time. So, there are 2 options are available for dividing image into zones. The first one is transforming the image into 2D and then do normal segmentation to get zones. But this method is not efficient or accurate, be-cause this method fall in losing many details. Since the body zone contains many poses, so transforming the image into 2D will produce only one view of the zone and the other views are hidden. As mentioned before the aps files contains 16 view of the image, so taking single view will produce image with less details. Also, hidden ob-jects might takes different shapes and have different poses. So, an object might be clear in one view while in other is not which reduce the accuracy of the algorithm.

The second option is dividing the image as it is – dividing all the 16 views of the image- into zones. So that each zone will be a 2D array that contains all 16 view of the zone. So for each image there will be a 17 zone space, and for each zone there will be a 16 2D view. Those views that are clear are set to true and views that are not clear at the zone are set to none. For example, zone 5 which represents the chest part of the body is visible in all angle view except the last eight view. And that is clear since the chest part is only clear in the front and the view around the front, while it is not visible in the back view and the views around the back. Also, zone umber 8 which represents the right hip is visible in all views except view number 3,4,5,6 and 7. This process produces 17 body zones that represents the body zones in the dataset so that we can feed the classifier each individual zone separately using the provided label each zone. Figure 7 shows the different views of zone number 1 which represents the right upper arm.



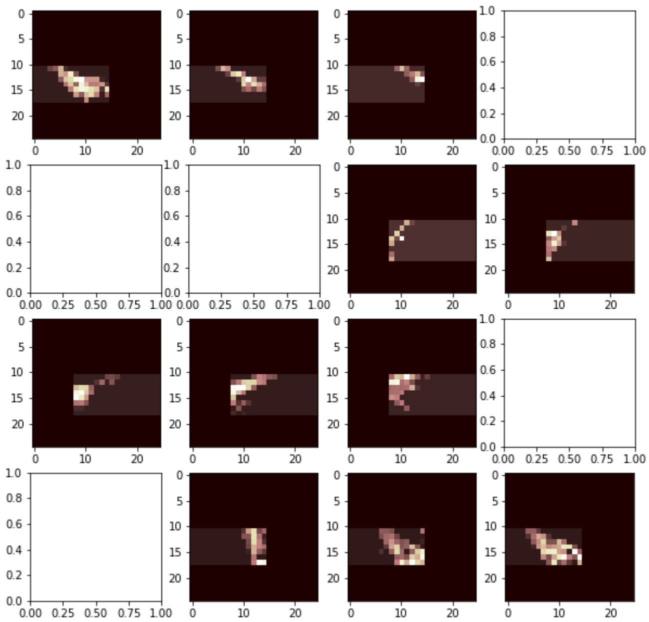


Figure 16*. This figure shows zone number 1 in the 16 different 2D views. Notice that those views where zone 1 is hidden set to none views (4, 5,6,12, and13). Where other view where zone number 1 is visible are represents into the cropped image.*

4.2.6 Cropped image Zero centred

Now we have for each aps image a 17 cropped zone images that can be handled individually as a separated image. Now the pixel mean of each zone is not zero centred, which make it difficult to the neural network to process the image. For example if a certain layer of the neural network needs to rotate the image and the mean of the pixel is not centred around zero, it will be a harder process for the network to perform. Zero centred process is used to transform the pixel centre to zero. It can be done by subtracting the pixel mean from each pixel, so that all pixels will be centred on zero not any other pixel. So if the mean of the pixels is a certain point rather than zero, all pixels have to be shifted to the zero point. Zero centring the image enhance the response of the convolution layer of the network; it increase the speed of the algorithm's learning process. Figure 11 shows the difference of normal data distribution and zero centred data distribution.

$$x = x - \sum_{i=0}^{n} p \qquad (3)$$

4.2.7 Cropped image normalization

Each zone might vary in the intensity level interval, although this variation in intensity should not affect the image analysis. Returning anything back to its normal state is the general meaning of normalization. However, normalization in image processing is rescaling the wide range of the image pixel intensity to a predefined range. So that the pixel intensity level of all image will be in the same range in other words uniform the distribution of pixel in all images.

By normalization we simply ignore the trivial change in the intensity levels, so that the model can learn the exact structure and image features like edges rather than learning the difference in intensity. This results in many advantages like speeding up the learning process of the model. In our problem we focus on detecting objects in the image through identifying the orientation and shape of the object through extracting features that represents the object. So if the same object is located in many images but in different intensity levels, normalization process eliminates the difference in the intensity and makes the two objects look the same.

The performance of many machine learning models is increased by normalizing and zero-centring the training and testing data feed to the model. Figure 11 shows the distribution of the data after the process of zero centre and then normalization. The figure shows how the distribution of the data has been centred around zero. And after normalization the range of distribution of the data has been reduced.

$$In = (I - Min)\frac{(newMax - newMin)}{(Max - Min)}newMin \qquad (4)$$

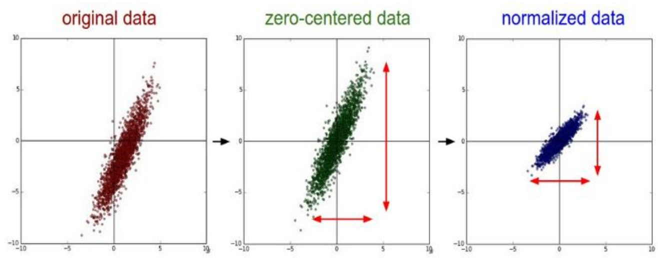


Figure 4*. This figure shows how the distribution of the data would change when it is zero centered and normalized. Notice that distribution of the data has been centred around zero. And after normalization the range of distribution of the data has been reduced.*
http://cs231n.github.io/neural-networks-2/

4.2.6 Save cropped zone

Once each image has been cropped into 17 zones and each zone is normalized and zero centred. Now each zone and its label are saved in the storage for further pro-cesses like training and testing. Each zone is name as object name followed by zone number and the label of the zone is retrieved from the csv file. The 10 gigabytes abs files are cropped and produce 140 gigabytes of cropped images.

4.2.7 Divide data into training and testing

Cropped images are then retrieved and divided into training and testing sets. The data set is split into 60% training data and 40% testing set.

4.2.8 Convolution neural network

The process of classifying an image into predefined classes used to be done by two main methods: feature extraction where some information that describe the image are extracted and then classification using a certain classification model is used to predict to which class the image belongs. Many different fea-tures like SIFT, Saliency, HOG, and other features can be extracted to describe provided image. Then the rule of the classifier is done by training the classifier with some images and their classes. There are many different classification model that can be adapted like decision tree, SVM, and other models.

However, the method of classification has proved its accuracy and effective-ness in solving some classification problems; it has been proved that this method has some drawbacks. If the extracted features do not describe the predefined classes, then the accuracy will be low. Also, when the provided data to be tested vary in the features, this way of classification will not perform well. For exam-ple, if we used a feature like SIFT and HOG to classify an image into dog or cat. Using the traditional method, if all the training data was of a specific kind of dogs like German shepherd and testing data contains images of another kind of small dogs like corgi dogs, the result of the classification will not be efficient. In our problem, the shape, size, and other features of concealed objects are not stable or static. Because we cannot predict the features of the concealed objects. That's why the need for a learning algorithm has to be adapted in our solution.

Artificial neural network is known by its ability to simulate the way that hu-man brain is learning through the neurons the construct the network. The convo-lution neural network is a special type of artificial neural network which is spe-cialised in image processing. The CNN has made a revelation in image pro-cessing and object detection algorithms. It has proved its efficiency in classifica-tion and detecting objects. And not like traditional way of object detection, CNN does not just depends on predefined feature to classify a creation image to a cer-tain class [6] [7].

Moreover, the CNN can extract huge number of features and learn from train-ing. That is because of the convolution layer in the hidden layers. Moreover, CNN can also perform very well in cases where image contains noise and fea-tures are not well known. There are different types of CNN with different struc-ture of neurons and different layers. Different hidden layers of the convolution neural network have different representation of the image features. The lower layers extract as many as general features. While the higher layers convert these general features to more specific features that represents the classes to which an image supposed to be classified. Two of the most known neural networks are AlexNet and ConvNet [7].

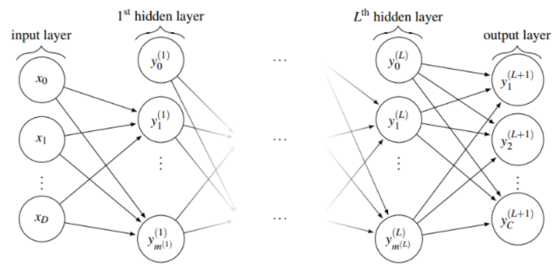


Figure 5. *This figure shows the structure graph of neural network layers and neurons. Input layer, hidden layers and the output layer [8]*

The CNN constructed of multilayers structure when one layer for the input, output layers and between them some hidden –convolution- layers. Each layer contains some neurons that assign a weight for each feature extracted by the neuron. Each neuron does not take the whole image, but instead it just focus only on a sub matrix of the image. Two main ways to assign the weights based on two techniques: the su-pervised method where these weights are predefined. And the second method is the unsupervised where the weights are selected randomly. These weights are reduced in the higher level where features that are more close to the classes are formed. In the training process, a method called back propagation is used to adjust the selected weights to best match the class.

AlexNet neural network is a special type of the neural networks that has been

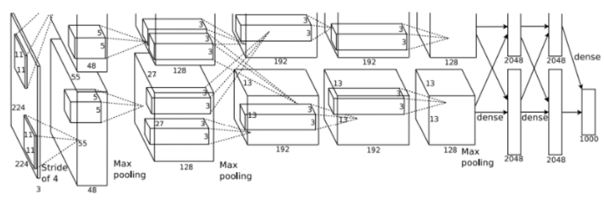


Figure 6. *This figure shows the internal structure of AlexNet CNN*

5. Summary

Airports security has become a serious issue that threats many people life. The incredible increase in the number of people who use the air for transportation make the security process at airports challenging. There have been many technologies that has been used for passengers screening problem. Previously this problem used to depend on the X-Ray backscatter technology. But the issues caused by the X-Ray waves like health problems and privacy has shifted the problem to the Millimetre wave technolo-gy.

The millimetre wave technology has no health side effects. And for privacy an automated target recognition algorithm has been applied to the MM wave technology where only the outline of the body is viewed, and the detected object is highlighted. With the image processing techniques, different methods have been adopted to detect concealed objects in MM wave images. In the related work we have targeted four of the most common ways. All these ways depend on the segmentation techniques, classification techniques, or combination of segmentation and classification. The segmentation method uses some edge de-tection algorithms to detect the objects. While classification method uses fea-tures extraction and the feed the extracted features to a classifier to classify the image.

References

[1] B. Elias, "Airport Body Scanners: The Role of Advanced Imaging Technology in Airline Passenger Screening," Congressional Research Service , 2012.

[2] O. Mart´ınez, L. Ferraz and X. Binefa , "Concealed object detection and segmentation over Millimetric Waves Images," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, Barcelona, 2010.

[3] S. o. Tapia, R. Molina and N. Blanca, "DETECTION AND LOCALIZATION OF OBJECTS IN PASSIVE MILLIMETER WAVE IMAGES," in 24th European Signal Processing Conference, Granada, 2016.

[4] S. Yeom and D. Su Lee , "Multi-level Segmentation for Concealed Object Detection with Multi-channel Passive Millimeter Wave Imaging," International Conference on IT Convergence and Security (ICITCS), Gyeongbuk, 2013.

[5] W. Ziye, Z. Ziran , L. Zheng, J. Yingkang, S. Zongjun and G. Jianping , "A Synthetic Targets Detection Method for Human Millimeter-wave Holographic Imaging System," in 2016 7th International Conference on Cloud Computing and Big Data, Beijing, 2016.

[6] K. Kiryong and H. Kim, "Classification of mixed-type defect patterns in wafer," IEEE, 2018.

[7] 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, vol. 13, p. 9, 2018.