Improving airport security screening: Introducing possible enhancements based on salience maps and the prevalence effect

Group 3

October 27, 2021

1 Problem Statement

International travel has become more accessible through the fast development of airports and its facilities. Furthermore, people can now fly to different destinations due to the progression of airplanes, different types of airlines and comfort classes. However, with more people circulating the globe, regulations and precautions are essential to protect one's safety and of others. Currently, every airport uses visual screening as a safety measure to detect dangerous and illegal objects. As this is the main instrument for detecting dangerous targets, the problem of detection errors and missing targets is severe.

1.1 Problem severity

The accuracy of visual screening tasks is vital for the safety of society as it detects unauthorised and dangerous items, for instance, a gun in a suitcase. Dangerous objects are rare targets and hence, are low prevalence items. Research shows that under visual searches, target rarity leads to an inaccurate performance in target detection. The effect of target prevalence on search performance in baggage checks is currently alarmingly low. The high error rate for low prevalence targets in screening tasks is severe. If similar effects would occur in socially important searches, the implications would have significant consequences.

1.2 Cause of missing low prevalence targets

To investigate the cause of missing low prevalence targets in visual screening tasks, an experiment was conducted that analysed the effects of target prevalence on search performance. The error rates as a function of the number of objects were measured. The results showed that errors increased dramatically and consistently as prevalence decreased. For example, a 10 percent prevalence rate returned a 16 percent error rate. With a 1 percent prevalence rate, the error rate increased to 30 percent. The errors were mainly due to "misses". The change in prevalence created a fourfold increase in error rates. Thus, the findings show that missing prevalence targets occur due to reaction time and the threshold for quitting when no targets are found. The threshold is constantly adjusted where observers slow down after making mistakes and speed up after success. For high prevalence targets, a fast "no" response causes mistakes. As a result, "no" reaction times were slower than "yes" reaction times. For low prevalence targets, "no" was given majority of the time. This drove down the quitting threshold. In other words, observers quit the search for low prevalence targets in less than the average time needed to detect a target.

1.3 Low prevalent target detection combined with other targets

Adding more prevalent targets with low prevalent targets in visual searches decreased the misses for common, high prevalence targets. Rare and very rare targets had a high rate of misses. In the second experiment, common (44 percent prevalence), rare (10 percent) and very rare (1 percent) targets were mixed for the search process. Only 11 percent of common targets were missed. However, 25 percent of rare and 52 percent of very rare targets were missed. This shows that the addition of high prevalence targets did not impact the detection rate of low prevalence targets.

1.4 Unfamiliarity with low prevalent targets

Low prevalence targets have a higher miss-error rate than high prevalence targets. When comparing the miss-error rate for 4.000 trials at a 1 percent prevalence rate, 40 targets were detected. However, low prevalence targets accumulated a high error rate of 41 percent. For high prevalence targets, the miss-error rate for 100 trials at a 34 percent prevalence rate found 34 targets and only had an 11 percent error rate. Therefore, the prevalence rate is critical for identifying targets and not only the targets presented.

1.5 Results in the real world

Visual detection of tools is performed through human signal detection tasks. The large increase of errors during the experiments suggests that this may be present in socially important search tasks as well. The results in the real world may be diminished by high quality visual detection training amongst professionals. Another key factor for identifying low prevalence targets is motivation. The conducted scoring system in the experiment can not replace the motivation of finding a low prevalence target, such as a gun, nor the motivation for the check-in line to move forward.

2 Cognitive task analysis for baggage screening

In order to investigate detection performance of illegal items during baggage screening at airports, a cognitive task analysis may be used. The cognitive task analysis identifies stages in human information processing. It also discovers the stage(s) where the problem of missing targets may come from. For the analysis, the 5 stages of cognitive information processing will be evaluated and the stage of the problem origin will be identified. The stages that will be analysed are:" 1) Sensory stage 2) Perceptual stage 3) Attention stage 4) Memory stage and 5) Decision- making stage".

2.1 Sensory stage

The screening starts by running the bag through the baggage scanner. The observer then analyses the image of the bag on the screen. At the sensory stage, the stimuli activates the observers sensory system. This enables them to see the image on the screen and absorb their surroundings.

2.2 Perceptual stage

Perceptual knowledge enables object recognition. During observation, concepts and their meanings are represented. Observers need to process what they are looking on the screen. The process involves drawing concepts and meanings based on previous experiences and from the process of calculating the causal factors of a set of observations that produced them.

2.3 Attention stage

Attention determines which information is ignored or processed. During screening, the observers divided and selective attention is activated. Divided attention is used to attend to two or more spatially separate locations simultaneously. Selective attention processes relevant sensory information and filters out irrelevant visual noise. The observer blocks out its surroundings and focuses on the screen to search for possible dangerous targets. Attention is necessary to process information, make decisions, and give the correct response. However, in complex tasks, demand for attention result from information overload, multi-tasking and making complex decisions. This may excel ones limited attention capacity. Furthermore, the cognitive burden for target searches is the interference between top-down standards for each target type. This may result in distorted target representations.

2.4 Memory stage

One of the activated memory stages during baggage screening is visual memory. The observer uses visual working memory where relevant information is retained by preventing the interference of irrelevant information. Illegal items in bags are often stored together with legal items. The observer has to identify the illegal items and disregard the legal items. One memory burden in random screening is short-term memory. It is the burden of recalling what area of the image has or has not been attended during random screening.

2.5 Decision-making stage

The third task involves decision making. The observer either approves the item and lets it trough, or, marks it as "suspicious" and sends it out for further analysis. The observer has to ultimately decide whether they are seeing a target or not.

2.6 Cognitive information processing stage of problem

Working memory capacity and attention control affect baggage screening performance. Working memory is needed to encode, retrieve and maintain templates. Maintaining the location and identity of a found target places a burden on memory and attention. As a result, this hinders the use of all available cognitive resources for the continued search.The problem of missing targets is also present at the decision-making stage. One problem is the decreasing search accuracy rate of rare targets.This is due to a criterion shift in decision-making. Observers become more biased to miss than to locate targets.

3 The center-bias and computational support systems

The center-bias is one of the human biases that can lead to the miss out of important objects, that are not aligned in the center of the image or screen that a person is viewing. In the following section, this particular bias is explained and addressed in more detail.

3.1 The center-bias in salience maps

The center-bias, that can be taken into account in salience maps, is a positional bias in, for example, the visual detection of objects, and tries to quantify it. Objects that are positioned more in the center of the image, are more likely to be detected / to receive attention. This bias can be recognized when mimicking the human attention mechanisms with a model like DeepGaze2.

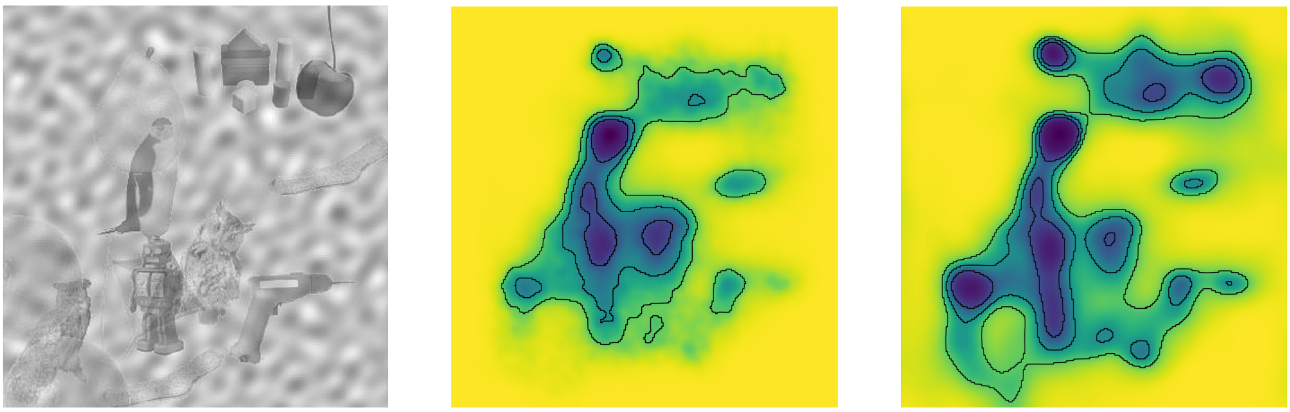

Figure 1: Salience Maps created by the DeepGaze2 model. a) Original Picture b) Salience Map with Center Bias c) Salience Map without Center Bias

The difference between a center-biased salience map and one without it, can be seen in 1: image b) and c). A center-bias causes the salience of the objects positioned around the center, to be more recognized. This is indicated by the darker color and black circles around the objects. However, objects that are further away from the center of the image, like in b), have a lot less attention, and thus a brighter color, on these objects. For a perfect observation, this should not be the case. All the objects should be equally treated, and thus, a salience map like in picture c) would be preferable. However, due to human biases, the salience map b), describing a scenario in which more central objects are getting more attention, is arguably more realistic. Such a model could also account for other known human biases. For example, a horizontal bias, or, a bias on written text or numbers.

3.2 Insights provided by the salience maps of Figure 1

Even though only one possible bias is taken into account, namely the center bias, assumptions of possible outcomes of the bias in a real-life task can be made. Assuming that the biased salience map b) is more likely to accurately describe human behaviour, the model, and thus, the human, is therefore more prone to miss certain objects that are further away of the center. Thus, detecting and quantifying such biases is a great step for creating less biased models, represented in map c). As a result, it can support humans in their tasks. However, it is important to point out that visual salience is not a one-size-fits-all solution to the problem of detecting malicious objects in airport security scans. A salience map that takes biases into account can support the human. It helps to locate the individuals focus attention. Nevertheless, this method does not take into account other factors such as shapes and visibility. A gun that looks like a teddy bear, or is (partially) covered by one, can still be overseen even if it receives a lot of attention. There are many more features to be considered and that can support the human, for example, anomaly detection to detect conspicuous shapes.

4 Research Experiment

In this section, an experimental setup is to be conducted, that measures the most likely causes of missing low-prevalence targets. One critical aspect is the application of eye-tracking technology. It can help figure out whether the misses originate from not looking at the locations of targets, so called search errors, or, whether the targets are looked at but not recognized (recognition/decision errors).

4.1 Experimental plan

When performing a visual search task, two main causes of missing targets can be identified. The first cause is search error and the the second is recognition/decision error. In order to investigate the relationship between error types and target identification during airport security screening, an experimental study needs to be conducted. One crucial point for investigation is the exploitation of eye-tracking technology. Eye-tracking devices measure eye position and eye movement. By tracking the observers gaze during the search process, the parts of the image which are deeply considered can be evaluated and consecutively, whether the area of interest is among them. As a result, it helps to understand the nature of errors. In the case of "misses", the two investigated errors are mutually exclusive (i.e. an individual is not subject to both throughout one attempt of target scanning, since, if the object is looked at the first place, checked by scanner data output, the nature of the error is determined). Based on this, the experiment aims to systematically minimize the biases that enable such flaws while singularly analyzing the impact and occurrence of the other one. For the investigation, it is necessary to handle the independent variables properly. In this case, they are the position and clarity of the object. When placed in an area where it is most likely to be seen, search error can be minimized. The maximum clarity of the object can minimize decision errors. However, the dependent variable is the score of the subject performing the scan.

4.2 Set up

The experiment is to be conducted at the airport. A pool of volunteers has to be chosen among people whose task is to regularly look at the monitor of the luggage scanner. The given experimental set up is used because of the vital importance of effectively reproducing the real life environment where such errors occur. Out of context experiments may affect vigilance and focus levels. They are amongst the main factors when studying low target prevalence errors in visual searches. Volunteers are to be seated in front of a 19-inch monitor with an average distance of 50 centimeters between the eyes and the screen. Every scanned luggage x-ray picture slowly slides from left to right while remaining entirely visible for approximately 3.5 seconds.

The observer's task is to locate dangerous object during baggage screening. if suspicion is raised, the luggage is stopped and send for further controls. If not, the item stays on the belt and goes through as usual.

Screen based eye tracker is used to keep track of participants gaze behaviour. The portable screen based eye tracker can be placed underneath the monitor, causing no disturbance for the individual performing the task. Such devices (like the *Tobii Pro/Fusion*) can have a sampling rate up until around 250 Hz. In this experimental scenario, it is strongly advised to have a sampling rate of at least 120 Hz, considering that the eye movements of interests are saccades (and eventual fixations) with an amplitude range that goes from 3° to around 25° max.

The experiment can be divided into two sections: In the first section luggage containing dangerous objects are presented to the volunteers (with real life like frequency) exclusively in areas where objects are most likely to be noticed (around the center). In section 1, the detection ability of the object must vary. From very familiar and clear objects (knifes or guns) orientated to be easily visible, to unique objects placed in unique ways (partially covered by other legal object). In sections 2, dangerous object are equally distributed all around the frame. For this part, it is recommended to have mostly medium or easy detectable targets.

Finally, after the experiment is extensively performed with a large number of volunteers, results can be compared with proper statistical tools to see the occurrence and the extent of the causes of missing low prevalence targets

5 Possible Solutions

This section presents a number of solutions for improving the rate of human error during visual search tasks at Airport Security. Although it is clear that the observers increased experience level in object recognition tasks could be very beneficial, several adjustments to the scanning process can be made to help detect suspicious objects. Hence, the following proposed solutions are system related. They aim to develop supplementary tools and equipment for humans to work with.

The first solution includes the use of a MRT like technique. It helps to visualize objects better by scanning the X-ray from different angles. This is especially useful for objects that might be covered by others. The auxiliary image from a different point of view could be displayed in case of doubt expressed by the employee or showed directly along with the standard from-above scan.

Additionally, improving the visualization of the scanned baggage to highlight objects that are further away from the center can be another solution. At the moment, displayed content is purely based on the object, having different colours for different materials. Adapting the brightness when the object is further away from the center could influence detection performance. Visualizing them differently based on their location on the screen can help tackle the center-bias problem and the errors that come with it.

Furthermore, a machine learning model could be used. The model should be trained to detect anomalies / suspicious objects and highlight them by segmentation or provide a warning when such an object is detected.

Lastly, to directly address the problem of low prevalence targets, deep fakes images could be randomly introduced into the daily scanning routine. Those could be artificial deep realistic pictures that present an expected suspicious target to be checked. Therefore, a physical fake luggage would no longer be needed for increasing the prevalence ratio of dangerous objects.