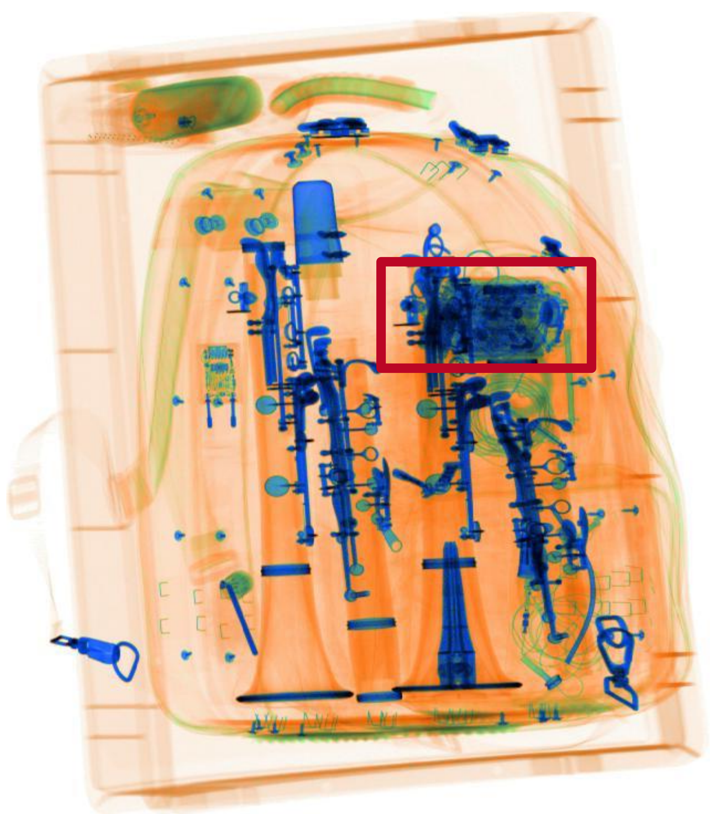
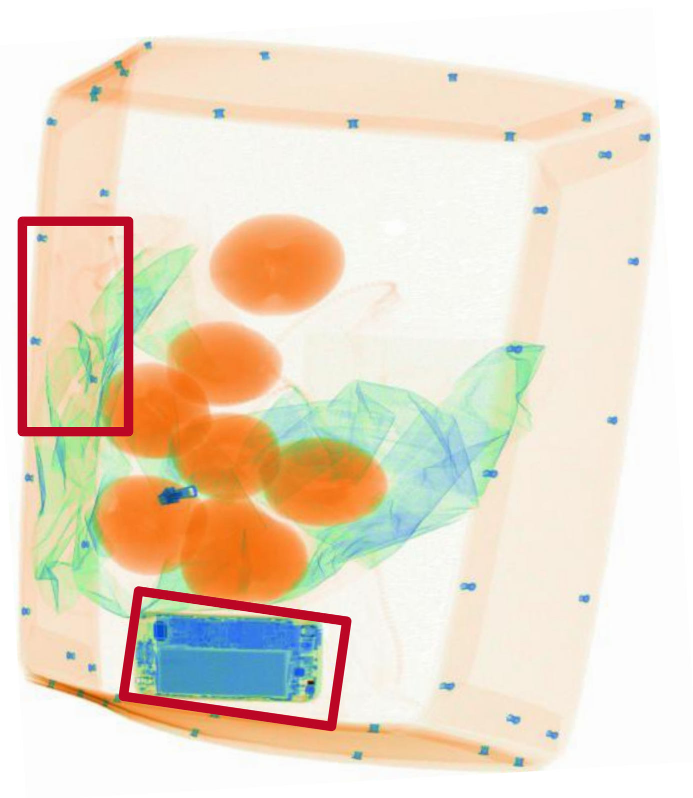
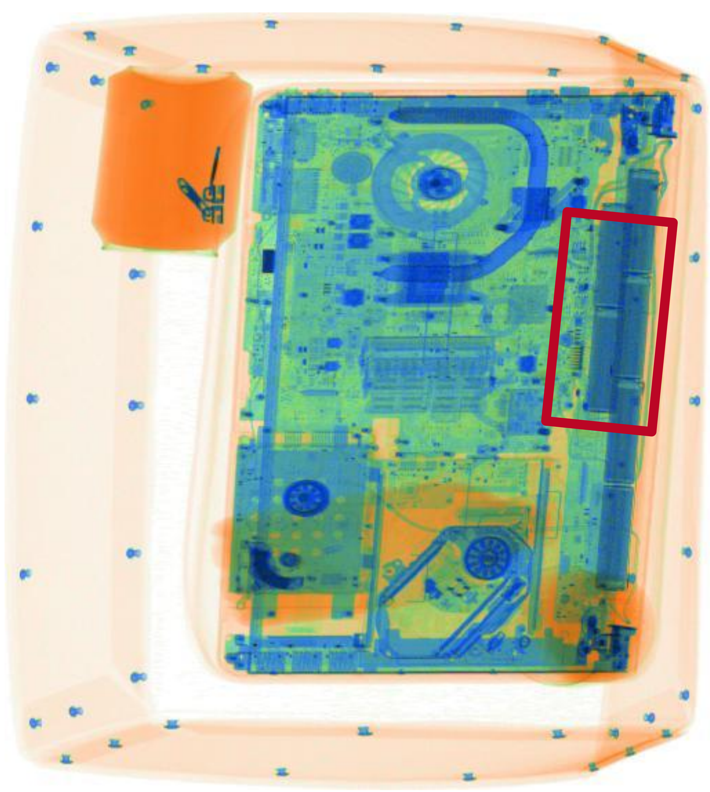
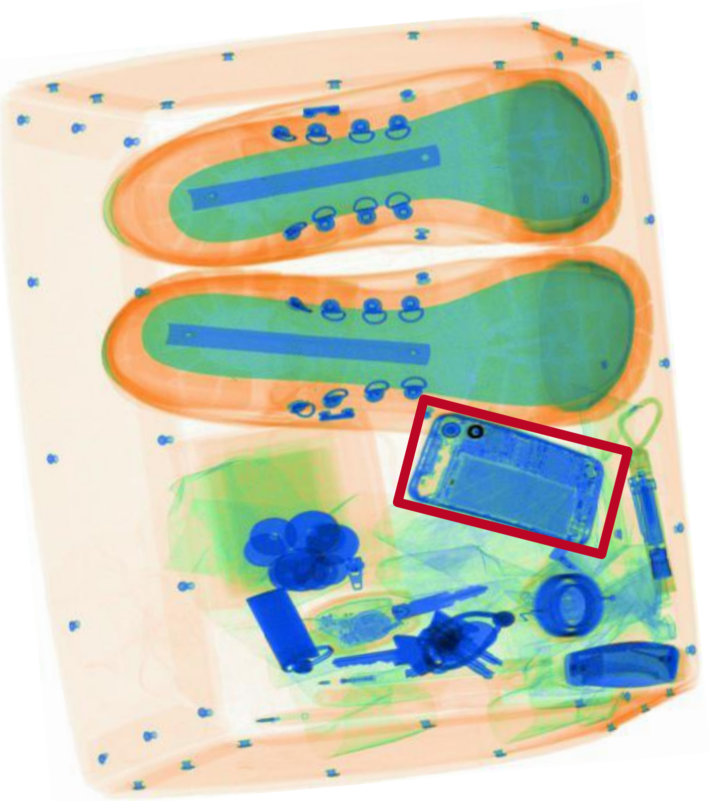
Detection of threat items: Airport security

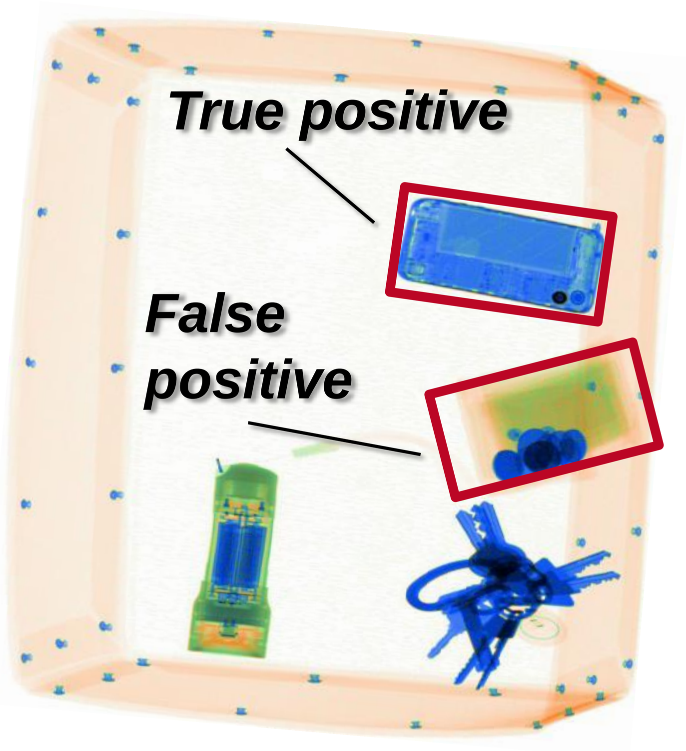
Mikkel Liisborg Hansen & Rasmus Jessen Aaskov, BSc. Students.
DTU Compute, Technical University of Denmark 2014.











Motivation

The X-ray inspection system is crucial for the security at airports. Our goal is to provide a tool for reducing the workload of the screeners and increase the overall security. In the following we introduce a method for object detection in X-ray scans. In our work we focus on detection on non-threat items such as smartphones, bottles, wallets and laptops.

Sampling

Given an object of interest, a retrieval of positive and negative images is performed, which is shown in Figure 1. From a database of 12,610 scans we have sampled 5,641 smartphones. Some of these will be used for training and others for testing.
Negative sampling is done by cropping images next to the object of interest to get the best distinction between the two classes.

Features

Among many available feature descriptors the combination of the following methods have proven to be effective for describing objects in X-ray scans.

- **HOG:** This method uses gradient images in the x and y direction to describe the magnitude and the direction of curvature. The method uses a contrast normalization to avoid variance in illumination. The spatial regions (known as cells) forms a histogram of the direction and magnitude sampled in 9 directions.
- **Color intensity histogram:** Sampling a histogram from each color channel.

The X-ray scans are given as RGB-images with colors corresponding to the materials (e.g blue represent metals), which we utilize by the color intensity histogram.
The HOG method has been the most popular method for previous related work due to its known high performance.

Classification

For a two-class supervised classification problem we use the linear Support Vector Machine (SVM). The SVM is a known method in image object recognition, because it is fast and can separate similar looking features for two different classes.
For class-specific retrieval we use a sliding window approach. Further more a cross-validation is performed to obtain the optimal variables e.g. HOG cell size, which can be seen in Figure 3.

Results

An overview of the performance of our method is shown in the ROC (see Figure 4). The differences of the true positive rate between the three implementations are clear. A choice of HOG cell size 26 gives the best overall prediction rate. Given a tolerance of 10% false positive rate, we obtain a true positive rate of about 68%, i.e. 4 out of 6 smartphones are correctly detected. If we only allow a 5% false positive rate, the true positive rate drops to approximately 55%.

A sample of X-ray scans are shown on the left hand side with bounding boxes indicating hits on smartphones. Note the false positive which has a similar shape as the a smartphone.

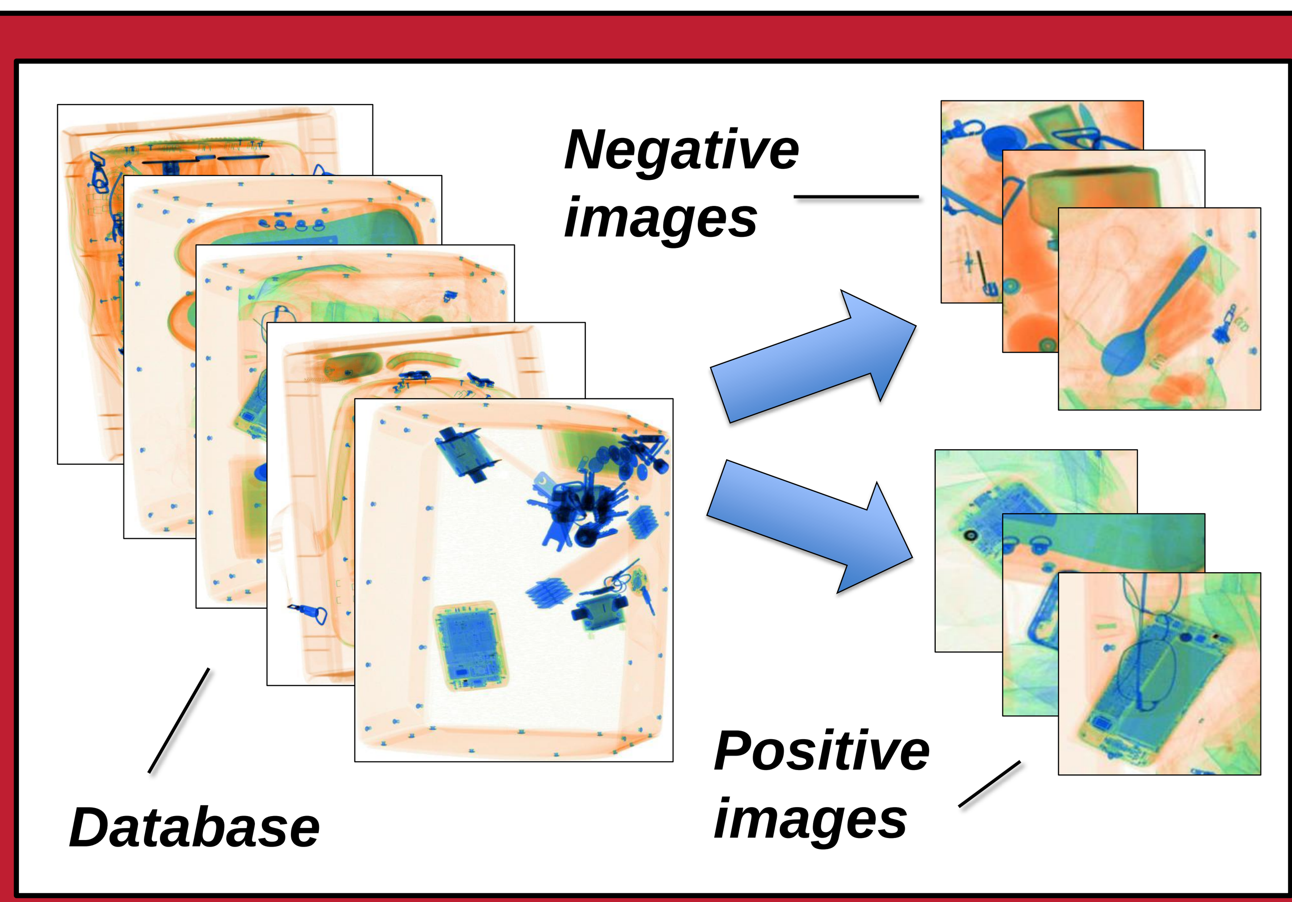


Figure 1: Sampled images form the foundation for feature descriptors.

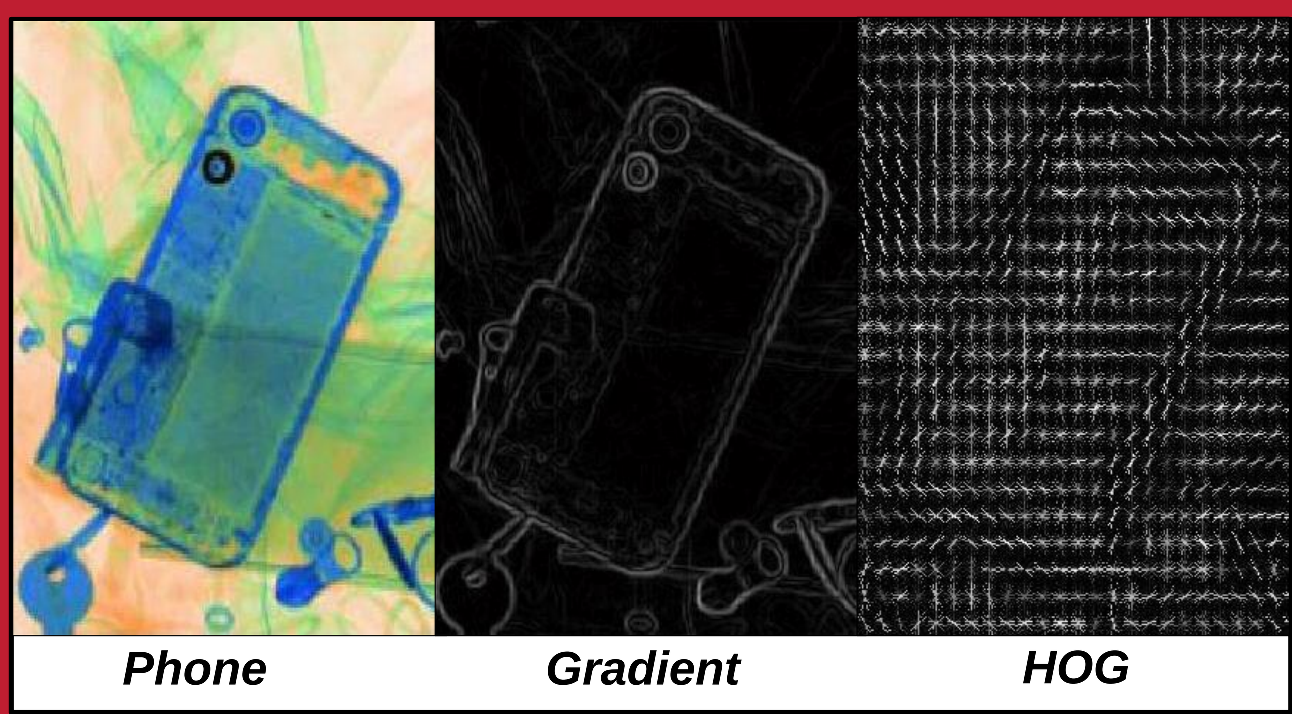
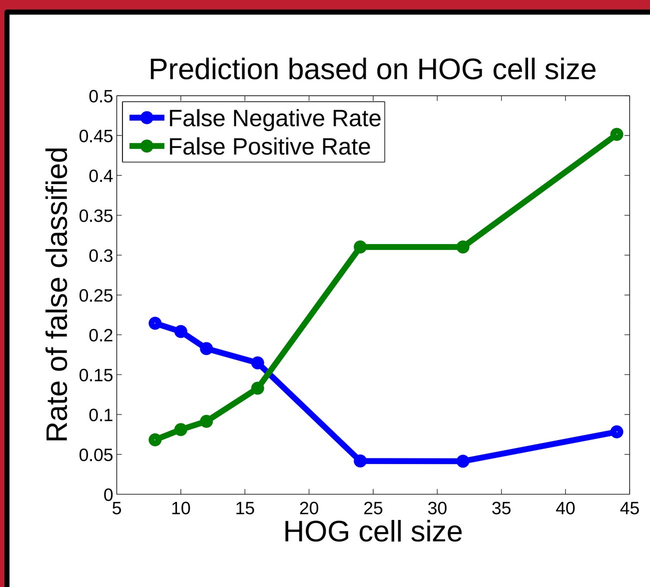


Figure 2: Gradient and HOG features shown on a smartphone.



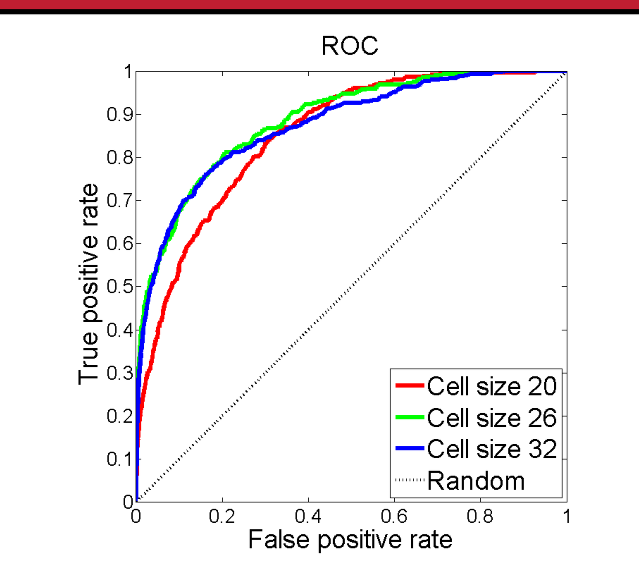


Figure 3: We see that we obtain the lowest false negative rate of smartphone with the hog cell size 26. However if we want to be even more certain about every classification of smart-phone we would need a lower hog cell size. But this will result in an decrease in the number of smartphones detected, hence a higher false negative.

Figure 4: The above Receiver Operating Characteristic (ROC) graphically illustrates the performance of our binary classifier. ROC analysis provides a tool to select possible optimal models and to discard suboptimal ones independently of the class distribution.

Table 1: Statistical results from untrained scans containing smartphones.

Cell size	PPV	TPR	F1-score
32	33,27%	94,62%	49,24%
26	37,16%	96,13%	53,60%
20	45,91%	95,62%	61,96%

The above results are based on a training set containing 1,705 scans. From Table 1 we see that the True Positive Rate (TPR) is best at cell size 26, as seen in Figure 3. However the best Positive Predictive Value (PPV) is found at cell size 20, which shows how often we classify an object correct. The harmonic mean of precision and sensitivity (F1-score) tells us that we gain a better overall performance choosing a lower cell size. An even lower cell size could be chosen but leads to a significant increase in computation time.