



Hidden Object Detection and Recognition in Passive Terahertz and Mid-wavelength Infrared

M. Kowalski¹ 

Received: 19 June 2019 / Accepted: 20 September 2019 / Published online: 8 November 2019
© The Author(s) 2019

Abstract

The study presents the comparison of detection and recognition of concealed objects covered with various types of clothing by using passive imagers operating in a terahertz (THz) range at 1.2 mm (250 GHz) and a mid-wavelength infrared (MWIR) at 3–6 μm (50–100 THz). During this study, large dataset of images presenting various items covered with various types of clothing has been collected. The detection and classification algorithms aimed to operate robustly at high processing speed across these two spectrums. Properties of both spectrums, theoretical limitations, performance of imagers and physical properties of fabrics in both spectral domains are described. The paper presents a comparison of two deep learning-based processing methods. The comparison of the original results of various experiments for the two spectrums is presented.

Keywords Hidden object detection · Object classification · Terahertz · Mid-infrared · Deep learning

1 Introduction

Non-intrusive detection and recognition of objects concealed under clothes remains challenging in terms of sensors and processing algorithms. Two spectra, namely thermal infrared (IR) and terahertz (THz), may provide images visualising objects placed on a human body covered with clothes. However, both spectra rely on different phenomena and provide images of different qualities.

The most useful property of imagers operating in terahertz range is the ability to see through clothing thanks to high transmission of terahertz waves through popular textiles. Thermal infrared imagers are able to detect small temperature differences on the object's surface. Both

✉ M. Kowalski
marcin.kowalski@wat.edu.pl

¹ Institute of Optoelectronics, Military University of Technology, Gen. S. Kaliskiego 2 Str., 00-908 Warsaw, Poland

properties of imagers can be used to detect metallic or non-metallic, potentially dangerous objects hidden under clothes.

Terahertz imagers, which are one of the most popular non-intrusive imagers for concealed object detection, provide images with low signal-to-noise ratio and low spatial resolution. The thermal infrared imager provides higher spatial resolution but they relies on thermal contrast which decreases due to thermalisation effect.

The paper presents a study on passive imaging of concealed objects with automatic detection and classification algorithm.

The study contains theoretical analysis of concealed object visualisation including relationship between spectral bands, the theory between heat transfer through clothing and between a hidden object and a human body. The paper briefly reports on values of transmittance of radiation through several textiles in both spectral bands. The selected deep learning methods are meant to combine real-time processing capability with high recognition rates. The study is concluded with results and analysis of proposed processing methods.

2 Related Works

Terahertz (THz) and millimetre-wave (MMW) are the most frequently explored spectra for detection of objects hidden under clothing with most of the commercial scanners operating in the MMW [1, 2]. However, mid-wavelength infrared is another promising spectral domain since it provides much more details than THz or MMW images [3]. Most of the systems for detection of concealed objects operate in MMW or THz spectra using various processing methods to automatically detect the object. Other related methods to reveal concealed items are based on THz spectroscopy [4–6] but their practical application in real-life is limited.

Various methods have been proposed for detection of concealed objects. Primarily, objects were detected by combining various feature descriptors with basic learning schemes. The most popular methods to create object representation include local binary pattern (LBP) [7, 8], Gabor Jet Descriptors [9, 10], histograms of Weber Linear Descriptor features [10] and histograms of oriented gradients (HOG) [11]. In recent years, machine learning (ML) methods, mostly based on convolutional neural networks (CNN), become increasingly popular. The state-of-the-art CNN architectures include the following: (1) single pass approaches that perform detection within single step (single shot multi-box detector (SSD) [12], you only look once (YOLO) [13]) and (2) region-based approaches that exploit a bounding box proposal mechanism prior to detection (faster regional-CNN (R-CNN) [14], region-based fully convolutional networks (R-FCN) [15], lightweight deep neural networks for real-time object detection (PVANET) [16], local R-CNN [17]).

Several approaches on hidden object detection have been proposed. Haworth et al. proposed several methods based on shape analysis. In [18], Active Shape Models (ASMs) have been combined with K-means to segment images into three regions: background, body and threats. Since the method was not accurate for body segmentation, another work [19] attempted to outperform using Gaussian mixture models. The results were better however unconnected body segments remained unresolved. Shen [20] combined various methods including anisotropic diffusion for denoising, mixture of Gaussian densities and isocontours for temperature modelling and segmentation. In [21], Martinez presented effective but time-consuming method based on non-local means (NLM) and iterative steering kernel regression (ISKR) for denoising and local binary fitting (LBF) for segmentation. Yeom proposed a

method using global and local segmentation, a Gaussian mixture model with parameters initialised using vector quantisation [22] and additional *k*-means clustering [23]. A method proposed by Agarwal et al. [24] uses a mean standard deviation–based segmentation technique. Gómez et al. [25] developed a two-step algorithm, based on denoising and mathematical morphology. Kumar et al. [26] used singular value decomposition and discrete wavelet transform for object detection and a convolutional neural network for classification.

Mohammadzade et al. [27] proposed an algorithm based on principal component analysis and a two-layer classification algorithm. Lopez-Tapia [28, 29] proposed a detection approach based on machine learning which learns from the spatial statistical information of the grey level. Liand et al. used mask conditional generative adversarial nets (CGANs) for real-time segmentation of weapons.

The paper concerns two deep learning methods, namely you only look once 3 (YOLO 3) and region-based fully convolutional network (R-FCN). The presented methods aim to detect and recognise (classify) hidden objects in THz and MWIR images.

3 Theory of Passive Concealed Objects Detection

The study concerns two spectra of complementary properties. Terahertz band is contained between frequencies from 0.1 to 10 THz (wavelengths from 3 mm to 30 μm) and is also commonly referred to as far-infrared [1–46]. This radiation is low-energetic and it does not ionise matter so it is not harmful to people. Infrared is radiation typically comprised in the range between 0.78 and 30 μm (430 THz down to 10 THz) [47], divided into four subbands: near, short-wavelength, mid-wavelength, long-wavelength and far-infrared.

Mid-wavelength infrared, also referred as thermal infrared, covers wavelengths from 3 to 8 μm (100 down to 37.5 THz). MWIR imagers are able to distinguish small changes of temperature [48] on an object's surface and therefore are useful in detection of objects covered with various types of fabrics [31, 49, 50].

Terahertz-based hidden object detection systems typically operate below 1 THz, due to high transmission through clothes in that range [51–54]. Ability to penetrate fabrics and other materials [55–57] with low water content is the basis for concealed object detection in the terahertz range.

It should be added that thermal infrared imagers, in contrary to imagers operating in THz spectrum, provide high-quality, high-resolution images with a higher spatial resolution. The spatial resolution in this context defines the smallest size of object that the imager is able to capture.

The transmittance depends on fabric type, basic weight and its thickness (number of layers) [58–60]. The transmittance through a fabric of terahertz radiation is significantly greater than of infrared and decreases with increasing frequency, number of layers and basic weight. Figure 1 presents the graphs of transmittance through selected clothes—shirts and a sweater made of cotton and polyester with different basic weights and leather jacket within the MWIR (3–6 μm) and THz (150 GHz–2.5 THz) ranges.

This study concerns passive imaging of concealed objects which relies on radiation emitted by objects. Passive cameras assign temperature differences to differences in the radiated energies in their spectral ranges per unit surface. The detection process relies on searching for temperature differences on the object's surface. Those differences should correspond to the differences of energy radiated in the whole radiation range. The larger differences may be classified as anomalies and hidden objects.

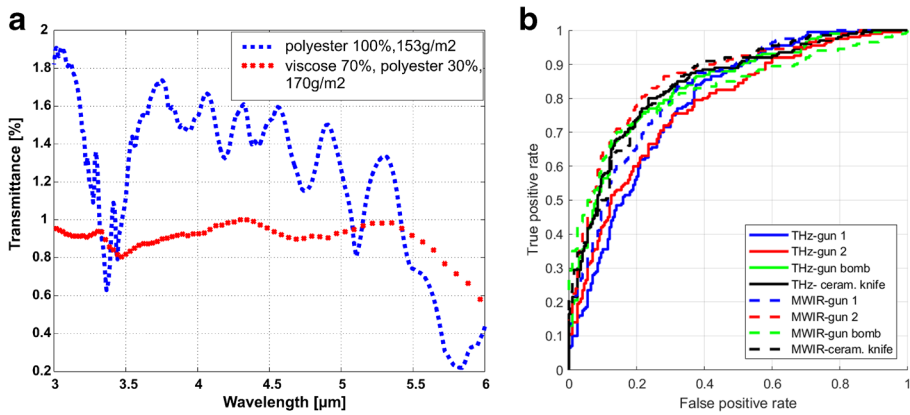


Fig. 1 Transmittance of **a** mid-wave infrared and **b** terahertz through different clothes

The ability of terahertz and infrared imagers to differentiate temperatures is the main measure for assessing their performance. The two parameters are minimum resolvable temperature difference (MRTD) and noise-equivalent temperature difference (NETD) with average values between 40 and 130 mK for cameras equipped with uncooled infrared detectors and between 20 and 70 mK for cooled infrared detectors that describe ability [50]. The value of NETD for THz imagers is lower and, typically in the range between 0.5 and 2 K.

The purpose of this study is to detect and classify potentially dangerous objects including guns and knives placed on the human body covered with the clothing. In the adopted measurement scenario, an item is placed on a human body with constant direct contact. The direct contact results in a heat energy transfer according to Fourier’s law [60, 61].

$$\vec{q} = -k\nabla T \tag{1}$$

where \vec{q} is the local heat flux density, k is the material’s conductivity (related to temperature) and ∇T is the temperature gradient. The Fourier’s law is often simplified and presented in one dimensional form of [61]

$$q_x = -k \frac{dT}{dx}. \tag{2}$$

Both bodies, according to the Planck’s law radiate the energy in a broad spectral range, including THz and IR frequencies.

The energy transfer between the object and the body is visualised in Fig. 2.

The total radiant emission is given with the following formula

$$\Phi_S = \tau_C \Phi_B + \Phi_C + \rho_C \Phi_A, \tag{3}$$

$$\Phi_W = \tau_C \Phi_H + \Phi_C + \rho_C \Phi_A, \tag{4}$$

where ρ_C is the reflection coefficient of a clothing fabric, τ_C its transmittance coefficient, Φ_B is the radiant exitance of the human skin, Φ_C is the radiant exitance of a clothing fabric, Φ_A is the irradiance of the clothing surface and Φ_H is the radiant exitance of a hidden object.

During the analysis of the energy transfer between the body, object and clothing, it was assumed that temperature of the external surface equalled the temperature of the object. The analysis does

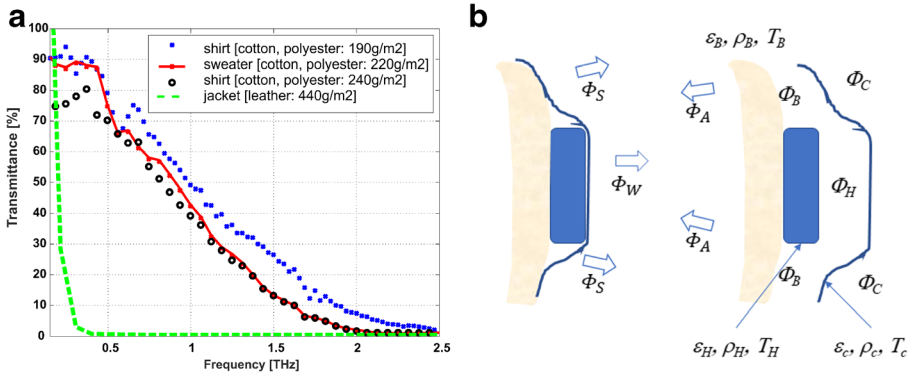


Fig. 2 a Thermal image of clothing surface with and without any object and b thermal radiation model

not take into account the transmittance coming from the human body that was transmitted, absorbed and re-emitted by the hidden object.

The spectral contrast $cp(\lambda, T)$ calculated for a given temperature (T) has been described by the following equation [13]:

$$cp(\lambda, T) = \frac{\varphi_B(\lambda, T) - \varphi_H(\lambda, T)}{\varphi_B(\lambda, T)}, \tag{5}$$

where $\varphi_B(\lambda, T)$ and $\varphi_H(\lambda, T)$ are the spectral radiant exitance of a human body and hidden item, respectively.

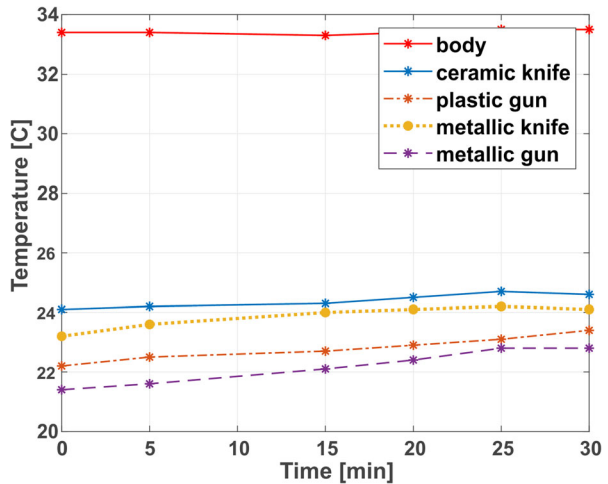
4 Experiment Protocol

The experiment concerned collecting images presenting the subject wearing various clothing with various objects hidden under clothes. In order to perform more detailed investigations, the set of objects to be concealed contained various dangerous objects as well as typical objects including wallet and mobile phone. During the experiments, images of a subject carrying various items were simultaneously acquired in both spectra. The dataset of collected images have been used to train and test the proposed algorithm.

The experiment has been divided into sessions lasting 30 min each. The aim of long-lasting measurement sessions was to collect images presenting various concealed objects with various contrasts and to assess the impact of decreasing contrast on detection and recognition.

The session duration is a result of initial experiments showing that concealed object used during this study reach the thermal equilibrium after 23–26 min as presented in Fig. 3. The time intervals apply to measurements performed indoors with constant ambient temperature. The duration of each

Fig. 3 Temperature values of concealed objects in time



measurement session has been adjusted to acquire possible changes that may result from the thermalisation process. During experiment, the object is heated by the body, and both reach thermal equilibrium. Measurement data (images, air temperature, humidity and pressure, values of the body and object temperatures) were collected every 5 min.

During each measurement session, a set of clothes and various metallic and non-metallic objects have been employed including a metal knife a plastic pistol, a ceramic knife and a subject mimicking dynamite. The set of objects is presented in Fig. 4. It has been taken into account that some of the objects may be mistakenly considered as different ones due to uniqueness of shapes. The set of items included also a leather wallet and a mobile phone.

All the measurement session were taken indoors in stand-off scenario with a subject standing in front of the imagers. The ambient temperature was controlled and constant (294 K, controlled by the air conditioning) and the relative humidity varied by no more than 3%.

Figures 5, 6 and 7 present sample MWIR and THz images collected during measurement sessions presenting a subject wearing various clothes with different objects hidden under clothing.

The experiments have been performed using MWIR and THz imagers with parameters provided in Table 1.

5 Method for Automatic Object Classification

This study involves two automatic methods for detecting the classification of concealed objects. Both methods are based on convolutional neural networks and are able to perform detection and classification tasks in a single architecture. Two deep learning-based methods, namely You Only Look Once 3 (YOLO3) and R-FCN, are described in the following subsections. The deep learning methods studied in this work follow two different approaches for detection and classification of objects in an image. YOLO3 is designed for fast object detection by eliminating the use a delegated region proposal network (RPN). R-FCN uses region proposal network to detect the object and softmax function for classification.

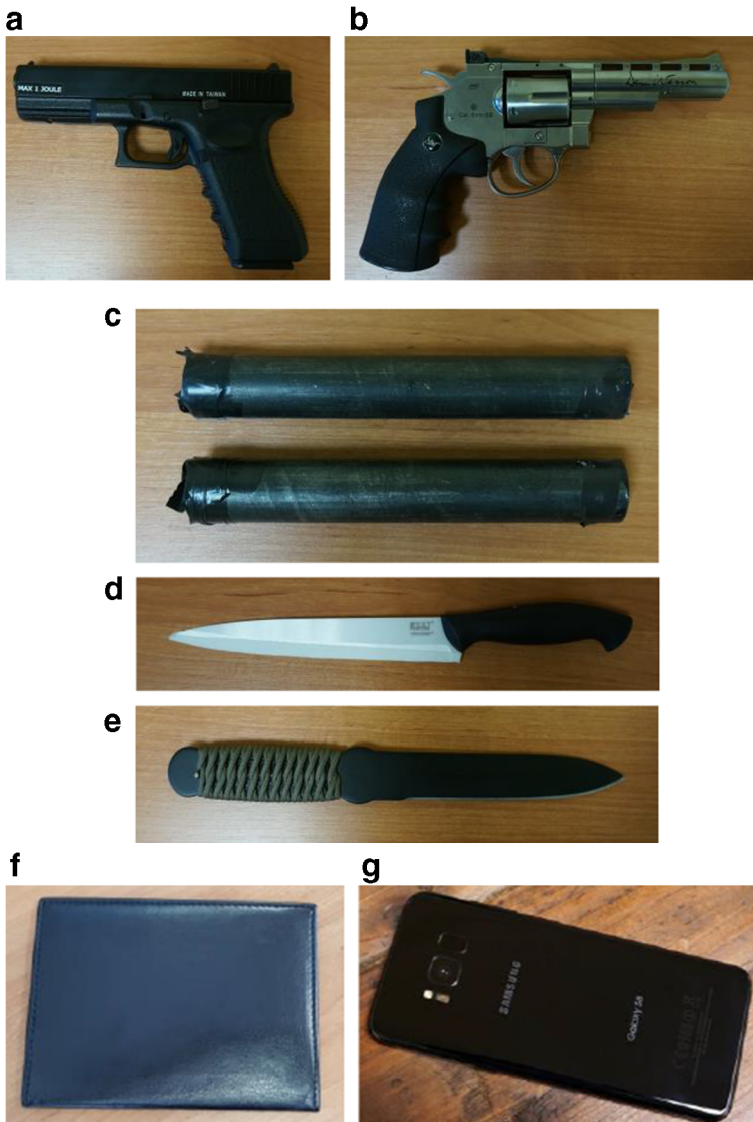


Fig. 4 Test objects used during the experiments: **a** plastic pistol, **b** metal pistol, **c** bombs, **d** ceramic knife, **e** metal knife, **f** leather wallet, **g** mobile phone

5.1 YOLO3

You only look once, also referred to as YOLO3, is known as a very fast object detection algorithm. The entire architecture contains 106 fully convolutional layers, where 53 layers come from the Darknet architecture trained on ImageNet. The architecture of YOLO3 is shown in Fig. 8.

The architecture is able to perform object detection and classification. The Darknet-53, a fully convolutional network with 53 layers, is used for feature extraction. The first detection is made by the 82nd layer. The objects' score for each bounding box are calculated using logistic

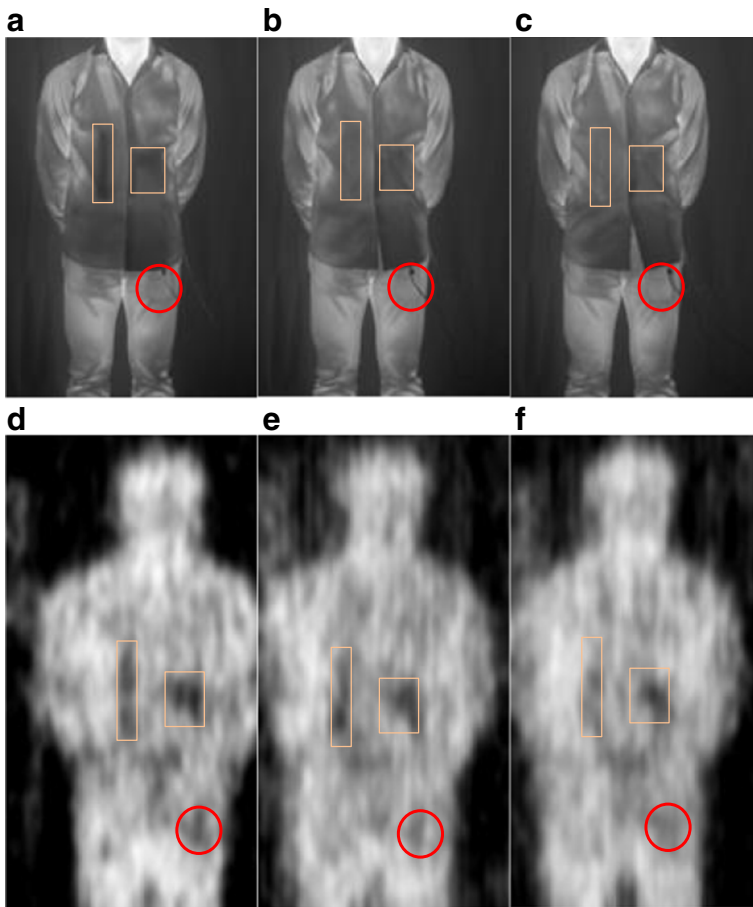


Fig. 5 Images presenting a subject wearing a thick cotton shirt, with bomb and a gun hidden under clothes. MWIR images at the beginning of the measurement (a), after 15 min (b) and after 30 min (c). THz images at the beginning of the measurement (d), after 15 min (e) and after 30 min (f)

regression at three scales. To determine the priors, YOLO3 applies k-means clustering algorithm. The 9 priors are grouped into 3 different groups according to their scale. Each group is assigned to a specific feature map.

The classification is done with independent logistic classifiers to calculate the likeliness of the input belongs to a specific label. Moreover, YOLO3 uses binary cross-entropy loss for each label which reduces the computation complexity.

5.2 R-FCN

R-FCN is based on the ResNet-101-based [38] framework. This architecture computes candidate regions by the fully convolutional region proposal network (RPN). The adopted region proposal and region classification approach are aimed to achieve detection and classification capability. It uses a bank of k^2 position-sensitive score maps for each category produced by last convolutional layer. The model is able to compute C number of object categories. The R-FCN architecture is presented in Fig. 9.

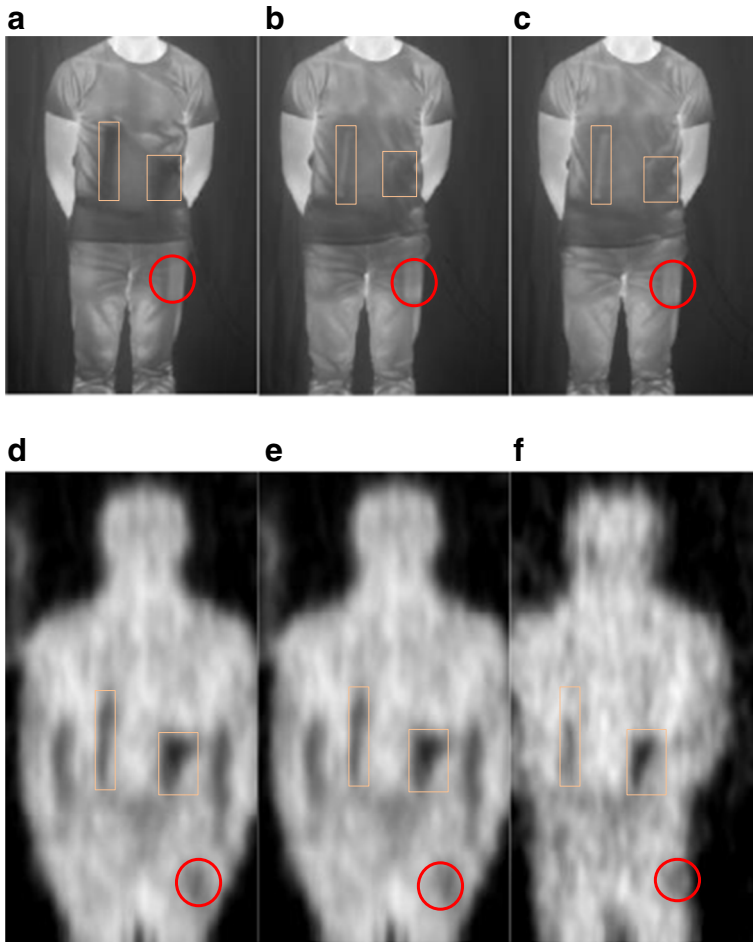


Fig. 6 Images presenting a subject wearing a cotton T-shirt, with metal bomb and a gun hidden under clothes. MWIR images at the beginning of the measurement (a), after 15 min (b) and after 30 min (c). THz images at the beginning of the measurement (d), after 15 min (e) and after 30 min (f)

R-FCN ends with a position-sensitive region-of-interest (RoI) pooling layer. This layer aggregates the outputs of the last convolutional layer and generates scores for each RoI.

6 Results

The results are clustered into two groups concerning detection and classification. During the 30-min acquisition sessions, the performance of object detection and object classification has been evaluated within 5-min intervals. At each specific minute of the experiment, both sensors acquired at least 250 images each. The dataset is composed of subsets containing images acquired during respective measurement sessions. The dataset has been split into train, test and validate subsets. Both models require large dataset to train upon, and the train-test-validation split ratio has been set to 75%, 10% and 15%, respectively. Data augmentation has been applied.

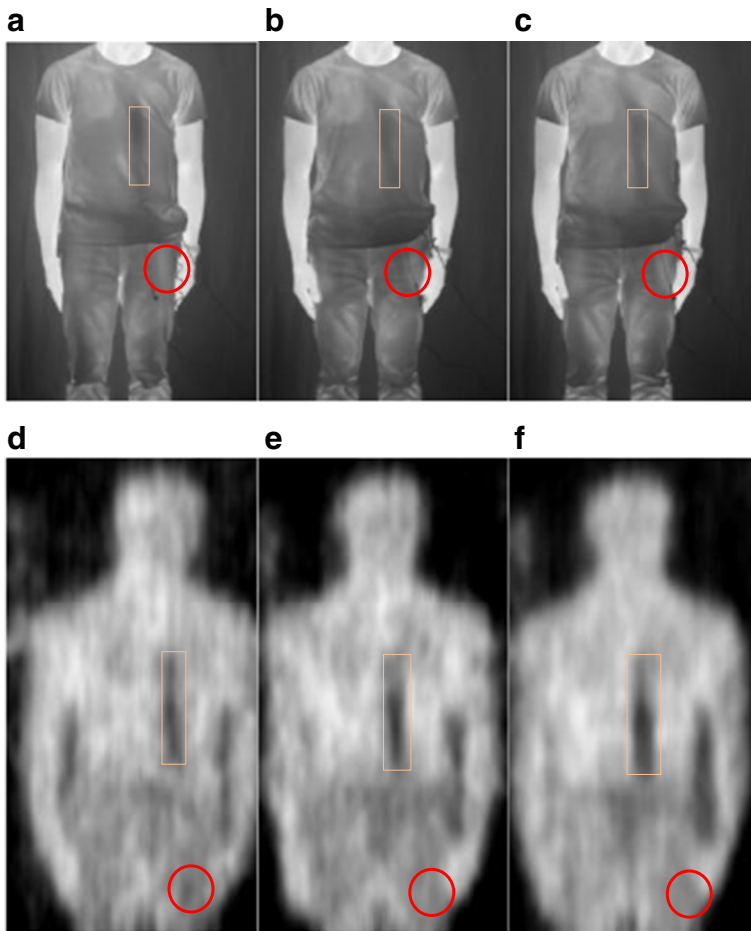


Fig. 7 Images presenting a subject wearing a cotton T-shirt, with ceramic knife hidden under clothes. MWIR images at the beginning of the measurement (**a**), after 15 min (**b**) and after 30 min (**c**). THz images at the beginning of the measurement (**d**), after 15 min (**e**) and after 30 min (**f**)

6.1 Detection

Performance curves of object detection task in various configurations are presented in Figs. 10, 11 and 12. During presented experiment, the subject was wearing a cotton shirt.

Table 1 Measurement equipment

	Parameters
MWIR imager	Focal plane array 640×512 pixels, MCT, cooled Spectral range: 2.0–5.7 μm Field of View: $25^\circ \times 21^\circ$ NETD: 18 mK at 300 K
THz imager	Passive sensing line, image 124×271 pixels, uncooled Frequency: 0.25 THz, ± 20 GHz (1.11 mm \div 1.3 mm) NETD: 1 K

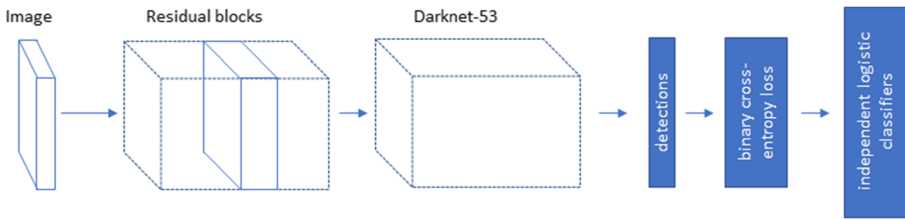


Fig. 8 Architecture of YOLO3

Our previous studies [60, 62] show that the ability to detect the concealed object with the infrared and terahertz cameras is decreasing for long-lasting experiments. The detection algorithm provided bounding box for every detected object in the image. The detection rate depends on the observed objects’ temperature, size, basic weight and type of clothing that the subject is wearing.

Comparison of the images indicated that in the case of the terahertz images, the contrast (the change in the normalised pixel intensity of the object) between the concealed object and the body decreases more rapidly in the case of MWIR images. Summary of the performance indicators is presented in Table 2.

The algorithms were trained to detect and classify certain objects with specific shapes. However, the non-dangerous items, which were not used to train the algorithm, were also detected. The detection rates of other objects (false acceptance rate—FAR) are considerably lower than desired items as presented in Table 2.

The results show that the presented detection method works better for MWIR images, but only at the beginning of the experiment. After several minutes, the detection rates for MWIR images are decreasing. The region-based method outperformed the single shot detector; however, the method is slower, allowing for detection at 11 and 7 frames per second for YOLO3 and R-FCN, respectively.

The maximum detection rates of 86% and 94% have been achieved for images registered at the beginning of the experiment with the metal gun hidden under T-shirt. This is the result of relatively high temperature difference between an object and subject’s body at the beginning of the acquisition session. Moreover, infrared images provide more accurate edges than respective THz images due to a higher spatial resolution. However, detection performance of the R-

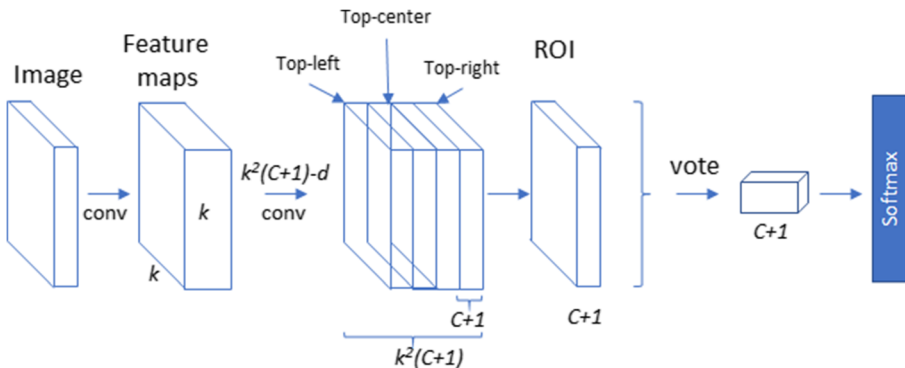


Fig. 9 Architecture of R-FCN

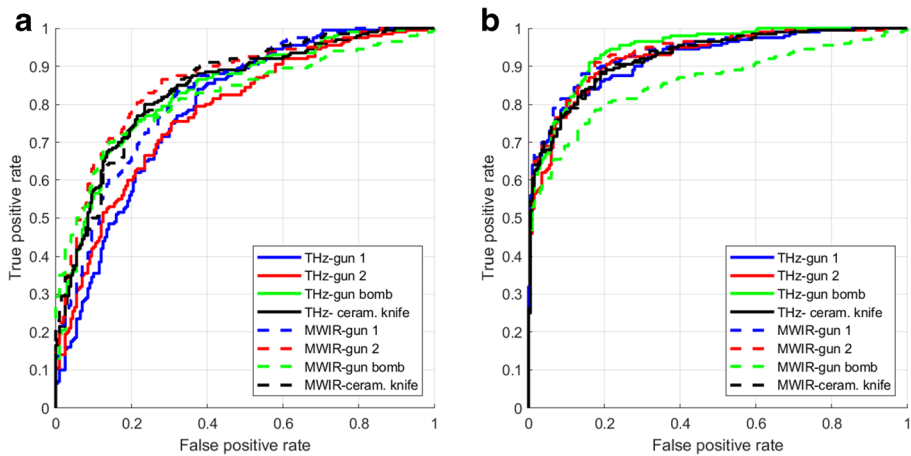


Fig. 10 Object detection performance. ROCs of **a** YOLO3 and **b** R-FCN for images collected at the beginning of the experiments in THz and MWIR domains

FCN for THz images was more stable, providing almost the same performance along the acquisition session.

6.2 Classification

The classification performance has been calculated only for the images with correctly detected objects of interest. Similar to the previous task, several experiments have been made. Performance curves of object detection task in various configurations are presented in Figs. 13, 14 and 15.

Both algorithms have been trained to classify following objects: guns, knives and objects mimicking dynamite. The mean performance of classification is lower than the detection.

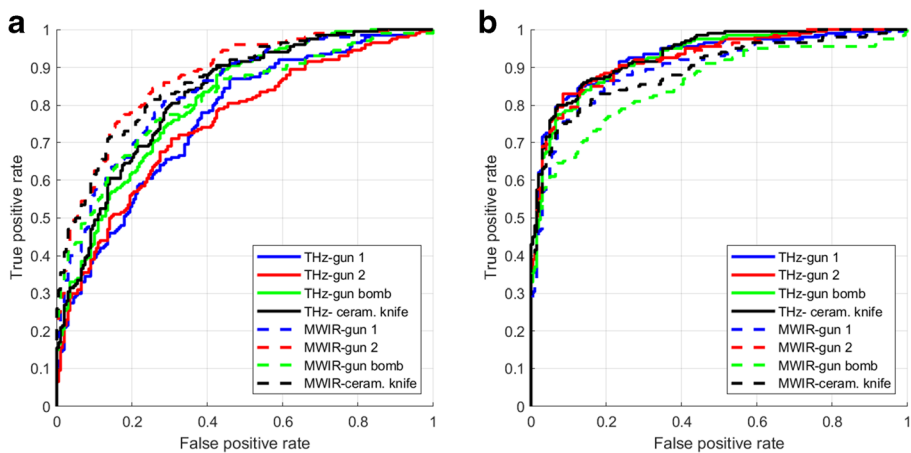


Fig. 11 Object detection performance. ROCs of **a** YOLO3 and **b** R-FCN for images collected after 15 min of the experiments in THz and MWIR domains

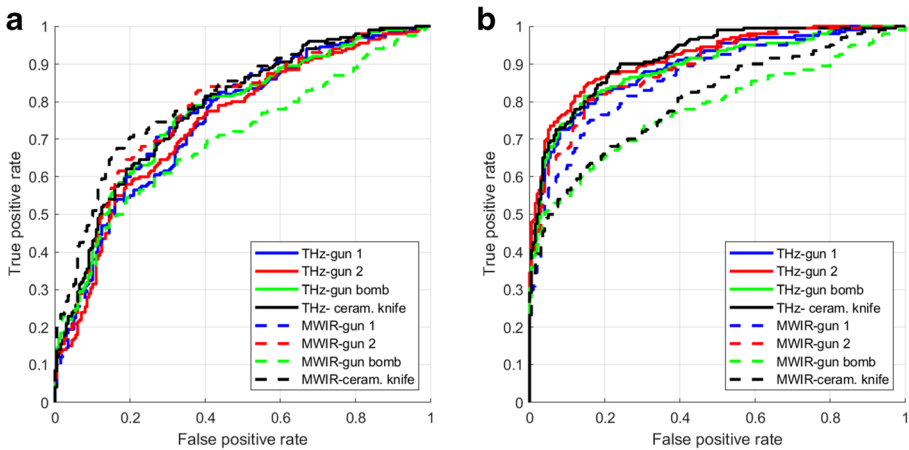


Fig. 12 Object detection performance. ROCs of **a** YOLO3 and **b** R-FCN for images collected after 30 min of the experiments in THz and MWIR domains

However, general trend is similar to detection task, where classification in MWIR domain outperformed THz range at the beginning of the experiment and showed decreasing rates along the experiment to finally perform worse than the THz domain at the end of acquisition sessions. Moreover, similar to detection task, R-FCN outperformed YOLO3. Classification rates of both methods for the THz images were slowly decreasing, providing almost the same performance along the acquisition session. Summary of the classification performance of presented algorithm is presented in Table 3.

Table 2 Performance of detection algorithm

Object type	AUC					
	Start		15 min		30 min	
	YOLO3	R-FCN	YOLO3	R-FCN	YOLO3	R-FCN
THZ						
Gun 1	0.78	0.92	0.74	0.92	0.74	0.89
Gun 2	0.77	0.92	0.74	0.92	0.76	0.91
Gun bomb	0.83	0.94	0.77	0.92	0.77	0.89
Cer. knife	0.83	0.92	0.78	0.93	0.73	0.91
Phone	0.07	0.05	0.07	0.05	0.06	0.03
Wallet	0.07	0.05	0.07	0.04	0.05	0.02
MWIR						
Gun 1	0.84	0.93	0.76	0.89	0.69	0.86
Gun 2	0.86	0.94	0.78	0.92	0.69	0.88
Gun bomb	0.82	0.86	0.70	0.85	0.68	0.78
Cer. knife	0.84	0.93	0.80	0.87	0.72	0.80
Phone	0.06	0.04	0.06	0.03	0.03	0.01
Wallet	0.05	0.04	0.05	0.02	0.02	0.01

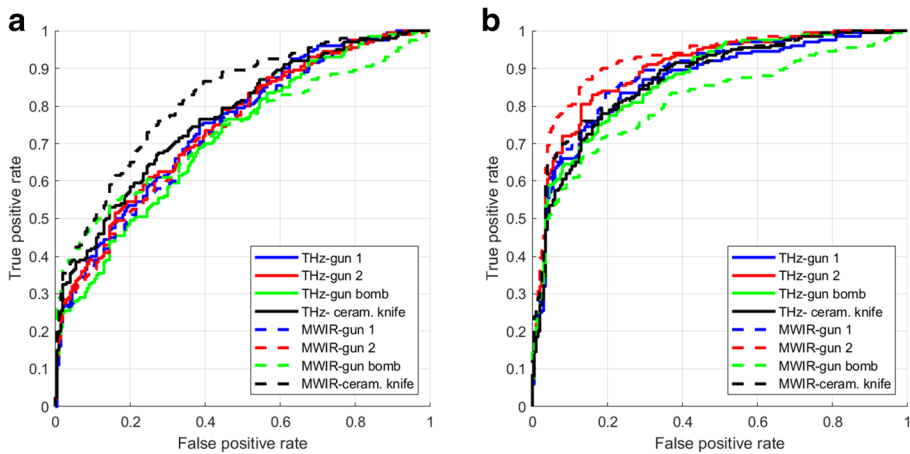


Fig. 13 Object classification performance. ROCs of (a) YOLO3, (b) R-FCN for images collected at the beginning of the experiments in THz and MWIR domains

Recognition rates vary depending on type of item and clothing. Detection and classification of concealed items depend on the number of factors, including the temperature difference between objects and human body, thickness, type of material and transmission of clothes. The general trend shows that large objects are easier to detect and classify than the small ones. Moreover, loose clothes may not perfectly adhere to the object and the heat transfer may not be perfect. This problem had particular importance for the imaging of objects within the infrared range where the transmission of clothes is very low and detection relies mostly on the heat transfer. When the contact between the object and the clothing is not uniform, the concealed object is not uniform and its shape may not be correctly classified.

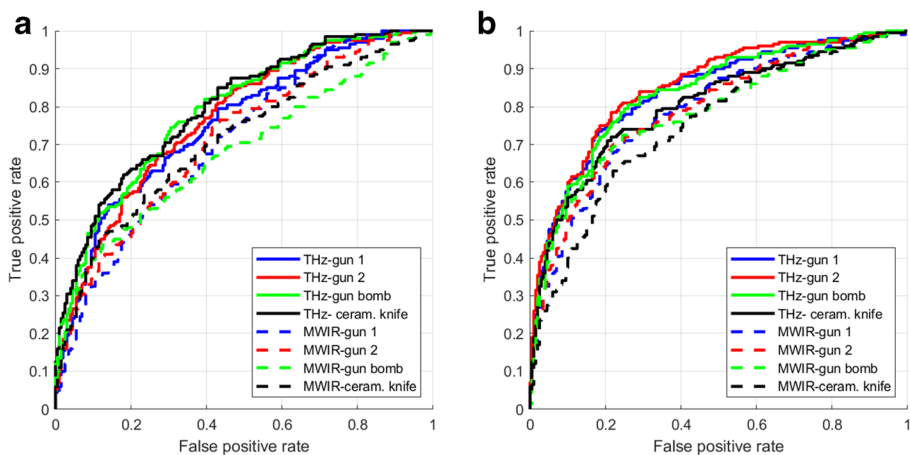


Fig. 14 Object classification performance. ROCs of **a** YOLO3 and **b** R-FCN for images collected after 15 min of the experiments in THz and MWIR domains

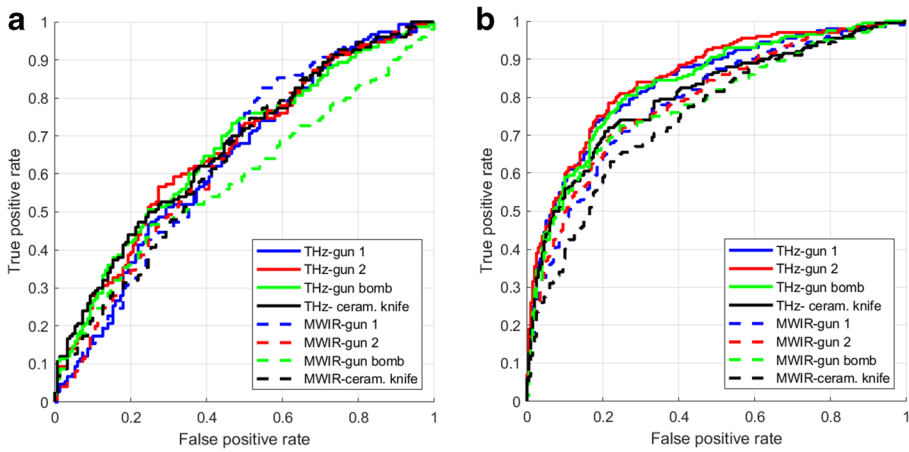


Fig. 15 Object classification performance. ROCs of **a** YOLO3 and **b** R-FCN for images collected after 30 min of the experiments in THz and MWIR domains

Since most of the known algorithms provide detection capabilities, only the result comparison is limited. Moreover, results obtained in different spectra and different types of imaging (active or passive) are not directly comparable.

A similar experiments performed with active THz imaging are presented in Table 4.

The presented methods have been implemented on the NVIDIA 1080-Ti graphical processing unit and are reported to process maximum 14 and 7 frames per second, for YOLO3 and R-FCN, respectively.

Table 3 Performance of recognition algorithm

Object type	AUC					
	Start		15 min		30 min	
	YOLO3	R-FCN	YOLO3	R-FCN	YOLO3	R-FCN
THZ						
Gun 1	0.69	0.86	0.63	0.87	0.63	0.83
Gun 2	0.74	0.89	0.64	0.88	0.66	0.85
Gun bomb	0.68	0.89	0.63	0.86	0.66	0.83
Cer. knife	0.72	0.86	0.67	0.88	0.66	0.80
Phone	0.02	0.01	0.01	0.00	0.01	0.00
Wallet	0.02	0.01	0.01	0.00	0.01	0.00
MWIR						
Gun 1	0.74	0.86	0.61	0.84	0.58	0.78
Gun 2	0.79	0.92	0.61	0.86	0.58	0.79
Gun bomb	0.76	0.81	0.57	0.79	0.59	0.78
Cer. knife	0.81	0.88	0.57	0.86	0.55	0.75
Phone	0.03	0.02	0.01	0.01	0.01	0.00
Wallet	0.02	0.01	0.01	0.00	0.01	0.00

Table 4 Comparison of detection performance

Method	Knife	Gun	Average
Yuan et al. [63]	65.24	69.54	58.45
Zhang et al. [64]	85.44	70.06	67.41
Yang et al. [65]	95.44	97.82	92.97
YOLO3	83.00	78.00	80.50
R-FCN	93.00	92.00	92.50

7 Summary

The process of a hidden object's detection and recognition has different physical sense in terahertz and infrared ranges and may be influenced by various factors. Terahertz cameras can visualise a hidden object mainly due to a non-zero transmission through clothing, whereas thermal cameras cannot utilise this property because the transmission rate through clothing is negligible. However, anomalies analysis of the temperature distribution on the clothing surface in the thermal infrared image we can reveal the object.

Generally, the classification task achieved lower performance than the detection in both spectra with the highest rates achieved by the region proposal network-based method. The experiments showed that the changes in the temperature values of the human body and the concealed object affect the detection ability of imagers operating in both the terahertz and infrared ranges. The compared methods, in some limited extent, are capable for real-time processing.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. L3 website, <https://www.sds.l3t.com/advancedimaging/provision-2.htm>, accessed 3/12/2018.
2. Smiths detection website, <https://www.smithsdetection.com/products/eqo/>, accessed 3/12/2018.
3. K. Ahi, "Developing terahertz imaging equation and enhancement of the resolution of terahertz images using deconvolution," *Proc. SPIE*, 9856, 2016.
4. K. Nakade, et al. Applications using high-Tc superconducting terahertz emitters, *Scientific Reports* 23178 (6), 2016.
5. N. Palka, R. Panowicz, M. Chalimoniuk, R. Beigang, "Non-destructive evaluation of puncture region in polyethylene composite by terahertz and X-ray radiation," *Composite part B Engineering* 92, 315–325 (2016).
6. K. Ahi, "Quality control and authentication of packaged integrated circuits using enhanced-spatial-resolution terahertz time-domain spectroscopy and imaging," *Optics and Lasers in Engineering* 104, 274–284, 2018.
7. Ojala, T., Pietikäinen, M., Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. *IEEE*.
8. Ojala, T., Pietikäinen, M., Harwood, D. (1996). A comparative study of texture measures with classification based on feature distributions. *Pattern Recogn.*, 29, 51–59.
9. Zou, J., Ji, Q., Nagy, G. (2007). A Comparative Study of Local Matching Approach for Face Recognition. *IEEE T. Image Process.*, 16, 2617–2628.
10. Chen, J., Shan, S., He, C., Zhao, G., Pietikainen, M., Chen, X., Gao, W. (2010). WLD: A Robust Local Image Descriptor. *IEEE T. Pattern Anal.*, 32, 1705–1720.

11. Dalal, N., Triggs, B. (2005). Histograms of oriented gradients for human detection. *IEEE Computer Vision and Pattern Recognition*.
12. W. L. et al. YOLO3: single shot multibox detector. *CoRR*, abs/1512.02325, 2015.
13. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: *CVPR*. (2016).
14. S. R. et al. Faster R-CNN: towards real-time object detection with region proposal networks. *CoRR*, abs/1506.01497, 2015.
15. J. Dai, Y. Li, K. He, and J. Sun. R-FCN: object detection via region-based fully convolutional networks. *CoRR*, abs/1605.06409, 2016.
16. K. Kim, Y. Cheon, S. Hong, B. Roh, and M. Park. PVANET: deep but lightweight neural networks for real-time object detection. *CoRR*, abs/1608.08021, 2016.
17. T. Vu, A. Osokin, and I. Laptev. Context-aware CNNs for person head detection. In *ICCV*, 2015.
18. Haworth, C.D., De Saint-Pern, Y., Clark, D., Trucco, E., Petillot, Y.R., 2007. Detection and tracking of multiple metallic objects in millimetre-wave images. *Int. J. Comput. Vis.* 71 (2), 183–196.
19. Haworth, C., Gonzalez, B., Tomsin, M., Appleby, R., Coward, P., Harvey, A., Lebart, K., Petillot, Y., Trucco, E., 2004. Image analysis for object detection in millimetre-wave images. In: *Passive Millimetre-Wave and Terahertz Imaging and Technology*, Vol. 5619, pp. 117–128.
20. Shen, X., Dietlein, C.R., Grossman, E., Popovic, Z., Meyer, F.G., 2008. Detection and segmentation of concealed objects in terahertz images. *IEEE Trans. Image Process.* 17 (12), 2465–2475.
21. Martínez, O., Ferraz, L., Binefa, X., Gómez, I., Dorronsoro, C., 2010. Concealed object detection and segmentation over millimetric waves images. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, pp. 31–37.
22. Yeom, S., Lee, D.-S., Jang, Y., Lee, M.-K., Jung, S.-W., 2012. Real-time concealed-object detection and recognition with passive millimeter wave imaging. *Opt. Express* 20 (9), 9371–9381.
23. Yeom, S., Lee, D.-S., Son, J.-Y., 2015. Shape feature analysis of concealed objects with passive millimeter wave imaging. *Progr. Electromag. Res. Lett.* 57, 131–137.
24. Agarwal, S., Bisht, A.S., Singh, D., Pathak, N.P., 2014. A novel neural network based image reconstruction model with scale and rotation invariance for target identification and classification for Active millimetre wave imaging. *J. Infrared Millim. Terahertz Waves* 35 (12), 1045–1067.
25. Liu, T., Chen, Z., Liu, S., Zhang, Z., Shu, J., 2016. Blind image restoration with sparse priori regularization for passive millimeter-wave images. *J. Vis. Commun. Image Represent.* 40, 58–66.
26. Kumar, B., Sharma, P., Upadhyay, R., Singh, D., Singh, K.P., 2016. Optimization of image processing techniques to detect and reconstruct the image of concealed blade for MMW imaging system. In: *IEEE International Geoscience and Remote Sensing Symposium*, pp. 76–79.
27. Mohammadzade, H., Ghoghoh, B., Faezi, S., Shabany, M., 2017. Critical object recognition in millimeter-wave images with robustness to rotation and scale. *J. Opt. Soc. Am. A* 34 (6), 846–855.
28. López-Tapia, S., Molina, R., Pérez de la Blanca, N., 2016. Detection and localization of objects in Passive Millimeter Wave Images. In: *24th European Signal Processing Conference*, pp. 2101–2105.
29. López-Tapia, S., Molina, R., Pérez de la Blanca, Using machine learning to detect and localize concealed objects in passive millimeter-wave images, *Engineering Applications of Artificial Intelligence*, 67 (2018) 81–90, 2018
30. P. K. Varshney, H. Chen, and R. M. Rao, On signal/image processing for concealed weapon detection from stand-off range, in *Proc. SPIE Optics and Photonics in Global Homeland Security*, T. T. Saito, Ed., 2005, vol. 5781, pp. 93–97.
31. M.C. Kemp. Millimetre wave and terahertz technology for the detection of concealed threats: a review, in *Proceedings of IEEE Infrared and Millimeter Waves*, pp. 647–648, 2007.
32. Jackson, J.E. (1991). *A User's Guide to Principal Components*. Wiley, New York.
33. Yu, H., Yang, J. (2001). A direct LDA algorithm for high-dimensional data – with application to face recognition. *Pattern Recogn.*, 34, 2067–2070.
34. Comon, P. (1994). Independent component analysis, A new concept?. *Signal Process.*, 36, 287–314.
35. Hermosilla, G., Ruiz-del-Solar, J., Verschae, R., Correa, M. (2009). Face Recognition using Thermal Infrared Images for Human-Robot Interaction Applications: A Comparative Study. *Robotics Symposium (LARS)*, 2009 6th Latin American
36. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L. (2008). SURF: Speeded up Robust Features. Springer, Berlin.
37. Lowe, D.G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vision*, 60, 91–110.
38. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, 2016.
39. Mateos, J., López, A., Vega, M., Molina, R., Katsaggelos, A., 2016. Multiframe blind deconvolution of passive millimeter wave images using variational Dirichlet blur kernel estimation. In: *IEEE International Conference on Image Processing*, pp. 2678–2682.
40. Yu, W., Chen, X., Wu, L., 2015. Segmentation of concealed objects in passive millimeterwave images based on the Gaussian mixture model. *J. Infrared Millim. Terahertz Waves* 36 (4), 400–421.

41. Dong Liang, Jiaying Pan, Yang Yu, Huiyu Zhou, “Concealed object segmentation in terahertz imaging via adversarial learning”, *Optik*, Volume 185, May 2019, Pages 1104–1114
42. M Kowalski, N Palka, M Piszczek, M Szustakowski, Hidden Object Detection System Based on Fusion of THz and VIS Images, *Acta Physica Polonica, A*. 124 (3), 2013.
43. John Johnson, Analysis of image forming systems, in Image Intensifier Symposium, AD 220160 (Warfare Electrical Engineering Department, U.S. Army Research and Development Laboratories, Ft. Belvoir, Va., 1958), pp. 244–273.
44. Y. S. Lee, *Principles of Terahertz Science and Technology*, New York: Springer, 2008.
45. X Zhang, X. Jingzhou, *Introduction to THz Wave Photonics*, New York: Springer, 2010.
46. P. H. Siegel, “Terahertz technology,” *IEEE Transactions Microw. Theory*, vol. 50, no. 3, pp. 910–928, Mar. 2002.
47. H. L. Hackforth, “*Infrared Radiation*,” Warsaw: WNT, 1963.
48. Y. Luo, W. Huang, “Attenuation of terahertz transmission through rain,” *Optoelectronics Lett.*, vol. 8, no. 4, pp. 310–313, Aug. 2012.
49. A. Rogalski, “History of infrared detectors,” *Opto-electronics Rev.*, vol. 20, no. 3, pp. 279–308, Sept. 2012.
50. A H Lettington, Q H Hong, An objective MRTD for discrete infrared imaging systems, *Meas. Sci. Technol.*, vol. 4, pp. 1106–1110, Apr. 1993.
51. R. Appleby, H.B. Wallace, “Standoff detection of weapons and contraband in the 100 GHz to 1 THz region,” *IEEE Transactions on antennas and propag.*, vol. 55, no. 11, pp. 2944–2956, Nov. 2007.
52. K. B. Cooper, R. J. Dengler, N. Lombart, B. Thomas, G. Chattopadhyay, P. H. Siegel, “THz Imaging Radar for Standoff Personnel Screening,” *IEEE Transactions on terahertz Sci. and Technol.*, vol. 1, no. 1, pp. 169–182, Sept. 2011.
53. R. Woodward, “Terahertz technology in global homeland security,” *Proc. SPIE*. 5781, 2005.
54. A. Rogalski, F. Sizov, “Terahertz detectors and focal plane arrays,” *Opto-electronics Rev.*, vol. 19, no. 3, pp. 346–404, Oct. 2011.
55. D. Mittleman, “*Sensing with Terahertz Radiation*,” Berlin: Springer Verlag, 2003.
56. D. Dragoman, M. Dragoman, “Terahertz fields and applications,” *Prog. Quantum Electronics.*, vol. 28, no. 1, pp. 1–66, Jan. 2004.
57. C. Jansen, S. Wietzke, O. Peters, M. Scheller, N. Vieweg, M. Salhi, N. Krumbholz, C. Jördens, T. Hochrein, M. Koch Terahertz imaging: applications and perspectives, *Appl. Optics.*, vol. 49, no. 19, pp. E48-E57, May 2010.
58. T. S. Hartwick, D. T. Hodges, D. H. Barker, F. B. Foote, “Far infrared imagery,” *Appl. Optics*, vol. 15, no. 8, pp. 1919–22, May 1976.
59. J. E. Bjarnason, T. L. J. Chan, A. W. M. Lee, M. A. Celis, E. R. Brown, “Millimeter-wave, terahertz, and mid-infrared transmission through common clothing,” *Appl. Phys. Lett.*, vol. 85, no. 4, pp. 519–521, Apr. 2004.
60. M Kowalski, M Kastek, Comparative studies of passive imaging in terahertz and mid-wavelength infrared ranges for object detection, *IEEE Transactions on Information Forensics and Security* 11 (9), 2028-2035, 2016.
61. J. H. Lienhard IV, J. H. Lienhard V, *A heat transfer textbook*, Cambridge: Phlogiston Press Cambridge, 2008.
62. M. Kowalski, M. Kastek, M. Walczakowski, N. Palka, M. Szustakowski (2015) Passive imaging of concealed objects in terahertz and long-wavelength infrared. *Appl. Optics*, vol. 54, No. 13, pp. 3826–3833, May 2015.
63. J. Yuan, C. Guo, A deep learning method for detection of dangerous equipment, Proceedings of the IEEE International Conference on Imaging Systems and Techniques (2018), pp. 159–164.
64. J. Zhang, W. Xing, M. Xing, G. Sun, Terahertz image detection with the improved faster region-based convolutional neural network, *Sensors*, 18 (7) (2018), p. 2327.
65. Yang et al. CNN with spatio-temporal information for fast suspicious object detection and recognition in THz security images, *Signal Processing Volume 160*, July 2019, Pages 202–214