



OPEN

Automated detection and classification of concealed objects using infrared thermography and convolutional neural networks

WeeLiam Khor^{1,2}, Yichen Kelly Chen³, Michael Roberts^{3,4} & Francesco Ciampa¹✉

This paper presents a study on the effectiveness of a convolutional neural network (CNN) in classifying infrared images for security scanning. Infrared thermography was explored as a non-invasive security scanner for stand-off and walk-through concealed object detection. Heat generated by human subjects radiates off the clothing surface, allowing detection by an infrared camera. However, infrared lacks in penetration capability compared to longer electromagnetic waves, leading to less obvious visuals on the clothing surface. ResNet-50 was used as the CNN model to automate the classification process of thermal images. The ImageNet database was used to pre-train the model, which was further fine-tuned using infrared images obtained from experiments. Four image pre-processing approaches were explored, i.e., raw infrared image, subject cropped region-of-interest (ROI) image, K-means, and Fuzzy-c clustered images. All these approaches were evaluated using the receiver operating characteristic curve on an internal holdout set, with an area-under-the-curve of 0.8923, 0.9256, 0.9485, and 0.9669 for the raw image, ROI cropped, K-means, and Fuzzy-c models, respectively. The CNN models trained using various image pre-processing approaches suggest that the prediction performance can be improved by the removal of non-decision relevant information and the visual highlighting of features.

The detection of concealed weapons and improvised explosive devices by terrorists has been a constant challenge for security operations. In populated venues such as airports and football stadia, it would not be practical for the security personnel to perform manual checks on all visitors. Ideally, a stand-off and walk-through scanner would determine if a threat exists, followed by the execution of a threat isolating or removal procedure, eliminating or minimizing the harm inflicted to civilians¹. Electromagnetic waves (EM waves) have often been exploited for non-contact object detection in human subjects¹⁻⁴. Some of the established EM waves used in security scanning include X-ray, millimetre-waves (MMW) and terahertz¹ (THz).

One of classic scanners is the X-ray detector⁵, which emits X-rays that are short in wavelength (10^{-7} to 10^{-9} m), high in frequency (3×10^{16} to 3×10^{19} Hz) and energy (124 keV to 145 eV). X-rays are the most penetrative among commercial EM scanners, giving full information on the cross section of the subject. However, the use of X-ray scanners is gradually being phased out due to the potential health implications when human subjects are exposed to ionizing radiation⁶⁻¹⁰. On the other hand, longer wavelength EM waves, such as MMW and THz have been effective in personnel security scanning applications. MMW has a frequency between 30 and 300 GHz with a wavelength between 1 and 10 mm, and can penetrate low visibility conditions, such as fog, to reflect off the subject¹¹⁻¹³. The penetration capability of MMW have led to the development of a commercial personnel threat detection system by QinetiQ, the SPO-NX¹⁴. Like MMW, THz enables good penetration through general clothing^{15,16}. THz has a frequency span between 100 GHz and 10 THz, and a wavelength between 3 mm and 30 μm . The penetrative capability of THz has led to the development of a commercial personnel scanning system

¹Department of Mechanical Engineering Sciences, University of Surrey, Guildford GU2 7XH, UK. ²Department of Technology, Design and Environment, Oxford Brookes University, Wheatley OX33 1HX, UK. ³Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK. ⁴Department of Medicine, University of Cambridge, Hills Road, Cambridge CB2 2QQ, UK. ✉email: f.ciampa@surrey.ac.uk

by Thruvision^{17–19}. Infrared (IR), when compared to MMW and terahertz, has a shorter wavelength (760 nm to 50 μm) and higher frequency (up to 214 THz). IR thermography captures thermal radiation emitted by objects in the form of image. There are two distinct forms of thermography: active thermography, where an IR illuminator (e.g., a flash or heater) illuminates the surface of the object, and passive thermography, where the IR radiation emitted from the surface of the object is directly sensed by the IR camera. Based on the black body radiation law, all objects above absolute zero degrees emit IR radiation, and variations of the radiation intensity are correlated to temperature gradients.

The potential use of IR technology in security scanning has been explored with various level of success^{3,20–27}. In cases of stand-off scanning for concealed object detection, the human subject acts as the heat source for passive emission sensing. The heat emitted from the body is absorbed by the concealed object and transferred to the clothing before emission off the surface of the top clothing layer. This generally results in a temperature gradient between the concealed object area and the clothing surface. However, emission signals detected from IR systems are less distinctive compared to the MMW and THz counterpart, as the wave penetration through clothing increases with wavelength²⁸. As a result, a low signal is expected when scanning through thicker layers of clothing using IR thermography^{21,29,30}. Despite these physical limitations, there are also inherent advantages of using IR systems for security scanning. IR has a shorter wavelength than MMW, which means it is possible to have a higher resolution when constructed as an image for visualisation, allowing detection from longer distances^{4,20,31}. It also enables higher-volume applications since multiple subjects can be visualised with larger fields-of-view. Additionally, images of faces captured using passive thermography are difficult to recognise, thus providing an additional layer of privacy²³.

In the context of concealed object scanning, most works have resorted to MMW or THz since they enable better penetration. Even though IR was less explored in this aspect, machine learning (ML)-based techniques have been successfully implemented with IR data, e.g., for non-destructive damage inspection and segmentation of moving objects^{32–37}. In an ideal concealed object detection system, the decision made upon scanning of the subject should be decisive and accurate. Due to the penetration limitation of IR systems, it could be difficult for an operator to make consistent decisions, particularly in cases of layered clothing.

The proposed approach in this paper to address this issue is by implementing convolutional neural networks (CNN) for concealed object detection and classification. Utilizing the transfer learning approach, a pre-trained CNN was fine-tuned using application-specific data, i.e., IR images of subjects with and without concealed objects. Transfer learning has been successfully implemented in many areas, such as biomedical imaging^{38–42}, non-destructive testing in buildings⁴³, agricultural pest and disease management^{44,45}, topology segmentation⁴⁶, image enhancement⁴⁷, pedestrian identification⁴⁸, and security scanning^{5,28}. The aim of this work is to explore the potential of using CNNs to predict the presence of a target object concealed underneath layered clothing using IR thermography. Raw IR images and pre-processed images were used to fine-tune a pre-trained ResNet-50 CNN model. This was followed by an evaluation of model performance using receiver operating characteristic (ROC) curves.

Methodology

Data acquisition

A FLIR A6750sc Mid Wavelength IR camera with a 3–5 μm waveband, < 20 mK thermal sensitivity (NETD), and a focal plane array size of 640 \times 512 pixels was used to capture thermal data. Examples of clothing and simulated concealed weapons used in data collection are shown in Fig. 1.

Image pre-processing

Four different pre-processing methods were investigated in this work: raw data from the IR camera, cropped region-of-interest (ROI), K-means clustered, and Fuzzy-c clustered images. This enabled us to investigate the effects of background removal, clustering, and information reduction on the trained CNN model for classification.

The ROI images were produced by cropping the subject from the raw image. Based on the generalised assumption of a 32 $^{\circ}\text{C}$ skin temperature, a threshold of 30 $^{\circ}\text{C}$ was used to crop the subject. This was applied on the raw image, where the first and last time the threshold temperature was encountered when reading temperature values from left to right of the image, and the first time when reading from top to bottom. This retains some of the exposed skin of the subject within the ROI image, removing some background information.

The K-means⁴⁹ and Fuzzy-c³² methods are unsupervised ML techniques used for information reduction. K-means is a well-known clustering method whereas Fuzzy-c is a less famous variation. Both approaches commence by establishing cluster centres to initiate the cluster affiliation. This is followed by computing membership scores based on distances between datapoints (pixels) and the cluster centres, which reveal the degree of association between each pixel and each cluster. These scores guide the adjustment of cluster centres, with more weights placed on pixels that exhibit stronger association. This iterative procedure continues until a defined stopping criterion is satisfied. The application of such information reduction technique on images will typically simplify the illustration of images. The primary distinction between these two methods lies in the fact that K-means employs hard clustering, restricting each data point to a singular cluster association (usually the nearest cluster centre)⁴⁹. Conversely, Fuzzy-c utilises a soft clustering approach, enabling each data point to be linked with all available clusters while assigning distinct degrees of importance (or weight)⁵⁰. The weighting (or membership score) is computed as,

$$\mu_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{D_{ij}}{D_{kj}} \right)^{\frac{2}{m-1}}}, 1 \leq i \leq C, 1 \leq j \leq N, \quad (1)$$



Figure 1. Photographs of simulated concealed weapons including (a) neoprene rubber block, (b) metal bearings in clay, and (c) fragments in money bag. Examples of outer clothing such as (d) layered windproof, (e) t-shirt, and (f) hoodie.

where μ_{ij} is the membership score of the j -th pixel (N pixels in total) for the i -th cluster (C clusters in total); D_{ij} is the distance (i.e., intensity difference) between the j -th pixel and the centre of the i -th cluster; and m is a parameter to control the extent of the fuzzy overlap. Updated cluster centres are subsequently computed as,

$$c_i = \frac{\sum_{j=1}^N \mu_{ij}^m x_j}{\sum_{j=1}^N \mu_{ij}^m}, 1 \leq i \leq C, \quad (2)$$

where c_i is the updated centre of the i -th cluster and x_j is the intensity of the j -th pixel. Furthermore, there is a disparity in the stopping criteria between K-means and Fuzzy-c. In K-means, the algorithm halts when cluster centres remain constant between consecutive iterations. In contrast, for Fuzzy-c, the algorithm continues its iterations until there is no further enhancement in the objective function. This objective function J is defined as

$$J = \sum_{i=1}^C \sum_{j=1}^N \mu_{ij}^m D_{ij}^2 \quad (3)$$

On general consensus, Fuzzy-c is considered more robust compared to K-means⁵¹. However, this could differ depending on application scenarios⁵². The thermal intensity data (apparent temperature) obtained from the IR camera was flattened into 1D before the pixel intensities enter the clustering algorithm. This allows the clustering algorithms to cluster the data relative to the apparent temperature in 1D, instead of area clustering in 2D. A total of 16 cluster centres were set to illustrate a total of 16 intensity bins in each image. Commercial numerical computing software, MATLAB⁵³, was used to process images and train models.

Convolutional neural networks, CNN

A ResNet-50 CNN model⁵⁴ was used to classify IR images. CNNs are a class of deep learning network architectures that can be applied to the classification of audio, time-series, signals, and most commonly for image classification. In the case of image classification, the input image is passed through a set of convolutional filters, each of which extract certain features from the images, e.g., brightness and edges Fig. 2.

The rectified linear unit (ReLU) serves as an activation function for the convolutional filter, and the pooling layer performs linear downsizing on the output, reducing the number of parameters that the network needs to learn. In the classification layers, the data is flattened into 1D before passing through a classification neural network, typically consisting of a fully connected network (also described as dense layer) and a SoftMax activation which normalises the output into probabilities. Typically, to train a CNN model from scratch, the general consensus requires that the size of the training dataset be at least one magnitude higher than the size of the test dataset to improve diversity and avoid overfitting^{55,56}. However, well-labelled IR datasets for subjects with concealed objects are rare and often not publicly shared. To tackle the small-dataset problem, a pre-trained ResNet-50 model was used. The model was pre-trained using the ImageNet dataset⁵⁷, which contained 1.4 million natural

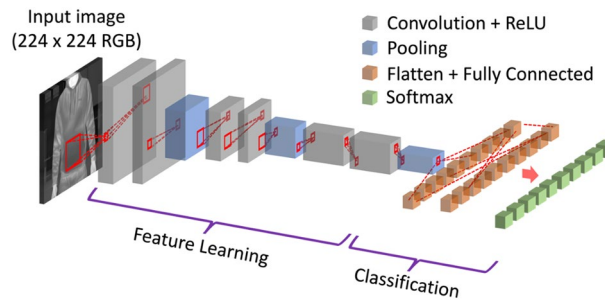


Figure 2. Architecture of the CNN image classification-based framework with multiple convolutional layers.

images with 1000 classes. In the transfer learning step, the fully connected layer was fine-tuned using a total of 900 labelled images (462 with and 438 without object). The underlying assumption in the transfer learning approach is that generic features extracted from an exceptionally large dataset are also present and informative in different datasets. This portability of learned generic features is a unique advantage of deep learning that makes itself useful in various domain tasks with small datasets⁵⁸. Using MATLAB's Deep Network Designer, the fully connected layer in the pre-trained model was modified to train using a weight learn rate factor and bias learn rate factor of 10. During training, argumentations were applied to the training images to improve the diversity of data and to avoid overfitting. Transformations were performed by applying reflection on the Y-axis, rotation by $\pm 45^\circ$, and rescaling by $\pm 25\%$. Internal validation was performed using 30% of randomly sampled training data to monitor the training process. Stochastic gradient descent was used as the optimiser with an initial learning rate of 0.0001, validation frequency of 5, maximum epoch of 20 and a mini batch size of 10. Validation accuracies of 98.95%, 97.19%, 94.07%, and 95.19% were achieved for the raw image, ROI, K-means, and Fuzzy C image models, respectively.

After training, a total of 200 images (equal split with and without object) that have not been exposed to the models (internal holdout/test-set) were pre-processed into the respective image type to evaluate the fine-tuned ResNet-50 models. When a cut-off is applied to the predicted probabilities, the comparison of the predicted and the true label will result in a confusion matrix consisting of true positives, false positives, false negatives, and true negatives Fig. 3. True positives and true negatives correspond to correct predictions, whilst false positives and false negatives are wrong predictions.

To evaluate the model performances without assuming a cut-off on the predicted probabilities, a receiver-operator characteristic (ROC) curve was used. This is a 2D diagram that is built using the true positive rate (TPR, also described as sensitivity or recall) and false positive rate (FPR, also described as false alarm rate) calculated using results obtained from the confusion matrix using each predicted probability as a tentative cut-off. These parameters are given by,

$$TPR = \frac{\Sigma TP}{\Sigma (TP + FN)}, \tag{4a}$$

		Actual class	
		Positive	Negative
Predicted class	Positive	TP	FP
	Negative	FN	TN

Figure 3. Confusion matrix for model evaluation.

$$FPR = \frac{\Sigma FP}{\Sigma(FP + TN)}. \quad (4b)$$

The TPR and FPR are metrics to evaluate the correct predictions of the model. Integrating the ROC curve, an area-under curve (AUC) can be obtained. The AUC provides a performance metric for the models, where 0.5 suggests no discrimination, 0.7–0.8 is considered acceptable, 0.8–0.9 is considered excellent, and > 0.9 is outstanding⁵⁹. The diagram in Fig. 4 illustrates the process of training and evaluation of the ResNet-50 models.

Results and discussion

The raw image illustrates the thermal intensity data captured by the camera data in 256 intensity values Fig. 5a. The ROI image crops the subject using a rectangular bounding box, removing excess background from the image Fig. 5b. An illustration of the K-means and Fuzzy-c post-processed images are shown in Fig. 5c,d respectively. The K-means and Fuzzy-c images looked like the ROI image, but with reduced information, as the images were illustrated with 16 grayscale bins instead of 256 bins. The cluster centre distribution in the K-means image is more even, making it look more similar to the ROI image. In contrast, in the Fuzzy-c image, clusters are more 'polar,' with distinct separation between cold and warmer areas.

The test-set ROC curves obtained from each model are shown in Fig. 6, revealing an AUC of 0.8923, 0.9256, 0.9485, and 0.9669 for the raw image model, ROI sectioned image, K-means and Fuzzy-c clustered image model, respectively. Based on AUCs, it is clear that pre-processed images tend to perform better than unprocessed data. Subject isolation by ROI cropping performed better than the raw image model, where some background not indicating the object presence was removed. K-means and Fuzzy-c both resulted in an improved background removal, where the background was entirely represented by a cluster. This offers a potential explanation to their superior model performance compared to the ROI cropped model. In the K-means model, the clustered image seemed more like the raw data where cluster transition is more gradual. In Fuzzy-c, the clustered image was more 'polar,' where the clothing showed a brighter cluster, and the object area was significantly darker. This could be a reasoning to the improved performance in the Fuzzy-c model compared to the K-means model, where the object area is more 'obvious,' allowing the CNN model to have more distinct features during the training process. Considering the improved performance of the post-processed image models, it is hypothesised that 2D clustering (1D clustering was used above), or advanced segmentation techniques could potentially further improve the performance of CNN models.

In the current study, concealed objects exhibited lower temperatures compared to the subjects. Consequently, pockets of lower apparent temperature, detected by the IR camera, indicated the presence of concealed objects on the surface of the clothing. Most of the clothing types used in this practice were mono-material, where the exposed clothing surface exhibited similar emissivity. However, in complex clothing, such as those comprising two or more materials with distinctly different emissivity, pockets or regions of low temperature can be 'artificially' created.

An image with an object concealed under complex clothing material is shown in Fig. 7. In the concealed object area, the torso area above the elbow exhibited a higher temperature reading compared to the area directly below the elbow, with a distinct border corresponding to the clothing design. The top part of the jacket is made of polyester (PET) with an approximate emissivity of 0.80, while the lower region of the jacket is made of cotton, with an emissivity of ~0.67⁶⁰. Therefore, the development of a robust model with the ability to identify concealed objects in complex clothing requires a large dataset with a variety of complex clothing materials for training.

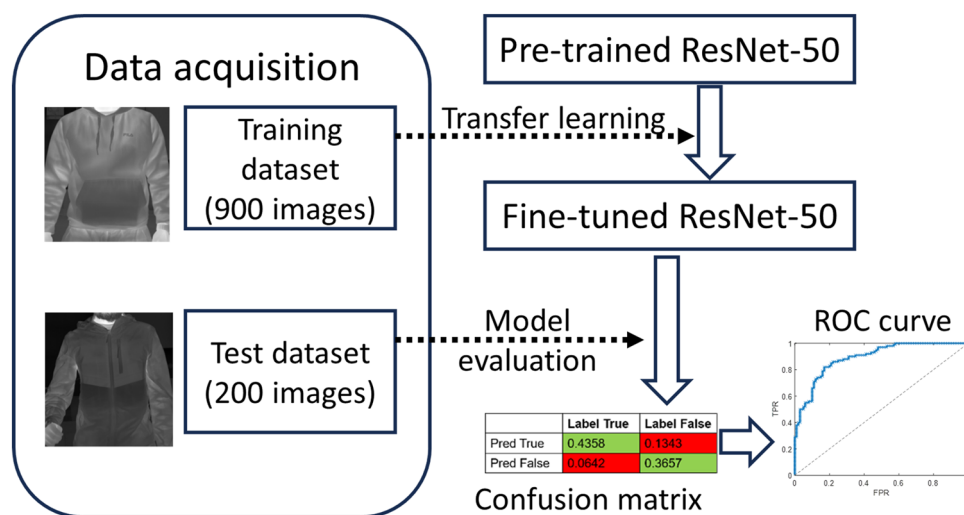


Figure 4. An illustration of the training and evaluation process of the ResNet-50 models for each of the different image type.

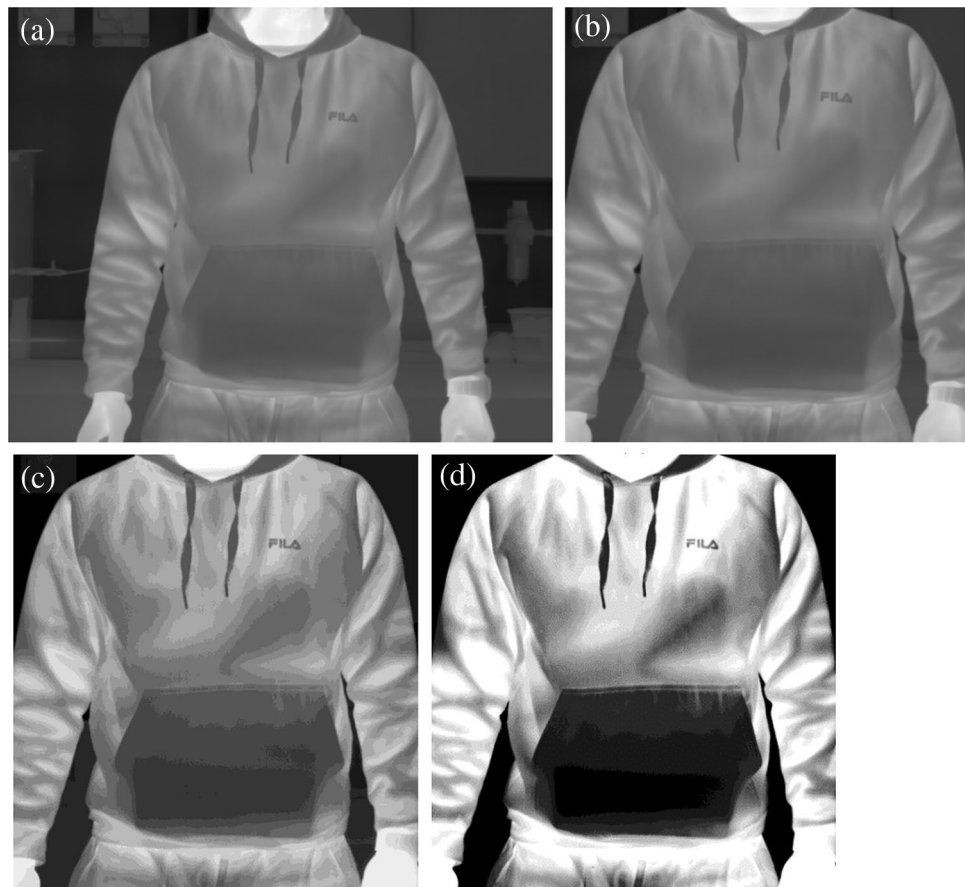


Figure 5. An example of the dataset image with a hidden rectangular box in the pocket, showing the (a) raw image, (b) ROI sectioned image, (c) K-means clustered image, and (d) Fuzzy-c clustered image.

Conclusions

Convolutional neural network, ResNet-50 was used to classify infrared images of subjects with concealed objects under clothing. The transfer learning approach was used to fine tune an ImageNet pre-trained model. Four models were developed and evaluated: a raw image model, ROI cropped model, K-means model, and Fuzzy-c model, which produced ascending improvement in model performance, exhibiting area-under-curve of 0.8923, 0.9256, 0.9485, and 0.9669, respectively. Pre-processing image data by removing non-decision relevant information from the images, such as background information, and visual highlighting of the object area effectively resulted in automated and enhanced model prediction performance. Regardless of image processing techniques, the application of transfer learning on a robust pre-trained network was shown to be effective on such small dataset problem.

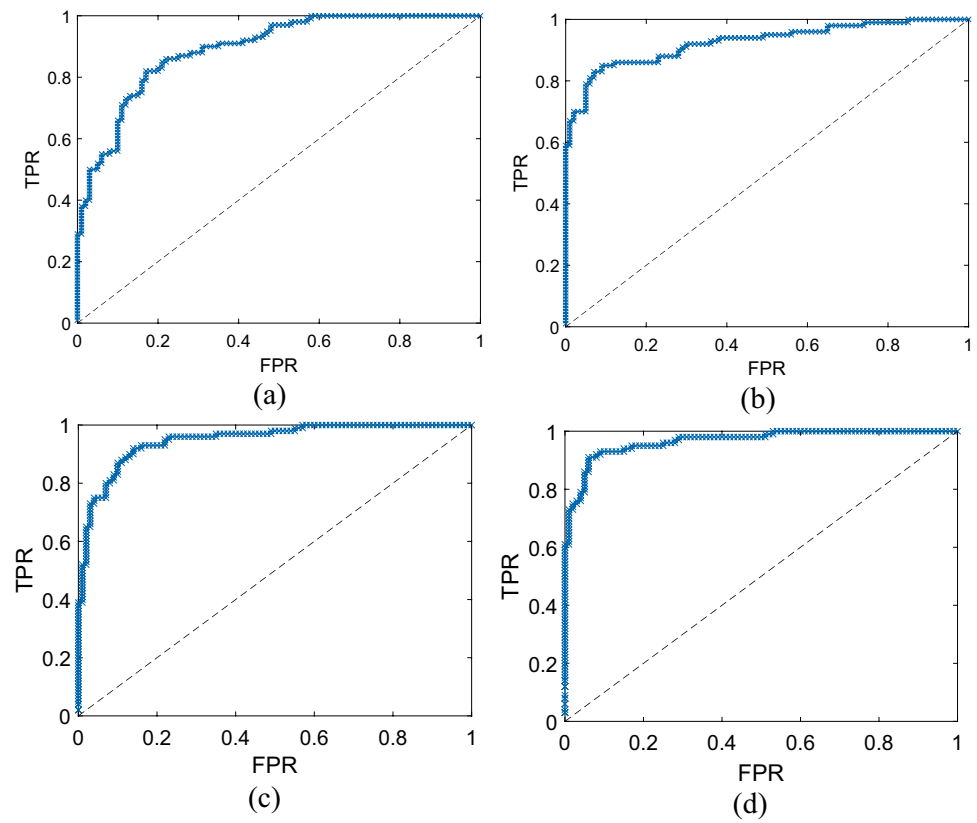


Figure 6. Receiver-operator Characteristics curves for the model trained using (a) raw images, and (b) ROI cropped images, (c) K-means clustered, and (d) Fuzzy-c clustered images.

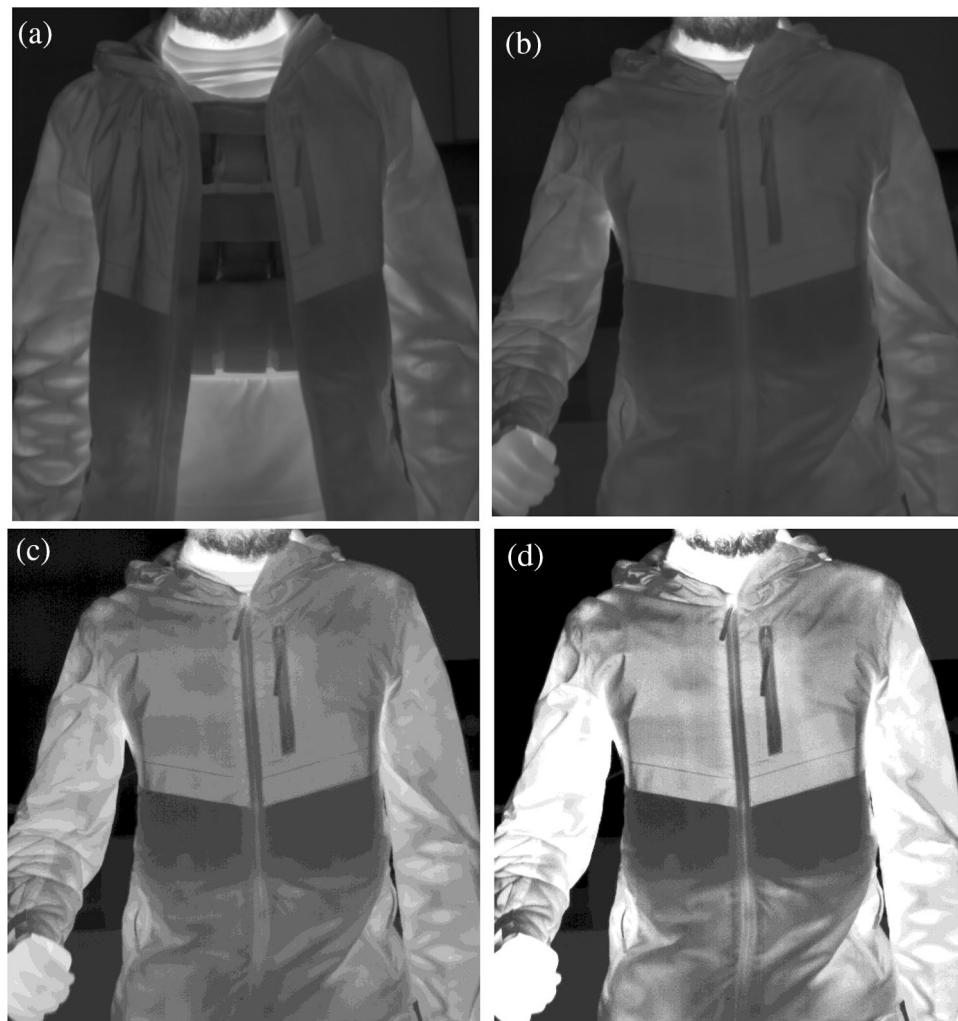


Figure 7. An example of a complex clothing with dissimilar materials with different emissivity. The ROI sectioned image with exposed object shown in (a), whilst the concealed object images were shown in (b) ROI sectioned, (c) K-means clustered, and (d) Fuzzy-c clustered.

Data availability

The data that support the findings of this study are available from the corresponding author, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the Defence Science and Technology Laboratory (DSTL), via the Defence and Security Accelerator (DASA).

Received: 18 October 2023; Accepted: 8 March 2024

Published online: 09 April 2024

References

1. National Research Council. *Existing and Potential Standoff Explosives Detection Techniques* (National Research Council, 2004). <https://doi.org/10.17226/10998>.
2. Cheng, Y., Wang, Y., Niu, Y. & Zhao, Z. Concealed object enhancement using multi-polarization information for passive millimeter and terahertz wave security screening. *Opt. Express* **28**, 6350–6366. <https://doi.org/10.1364/OE.384029> (2020).
3. Kasteck, M., Kowalski, M., Polakowski, H., Lagueux, P. & Gagnon, M.-A. *Passive signatures concealed objects recorded by multispectral and hyperspectral systems in visible, infrared and terahertz range*. Vol. 9082 SID (SPIE, 2014).
4. Kukutsu, N. & Kado, Y. Overview of millimeter and terahertz wave application research. *NTT Tech. Rev.* **7**, 6 (2009).
5. Akçay, S., Kundegorski, M. E., Devereux, M. & Breckon, T. P. in *2016 IEEE International Conference on Image Processing (ICIP)*. 1057–1061.
6. Barth, H. D., Launey, M. E., MacDowell, A. A., Ager, J. W. & Ritchie, R. O. On the effect of X-ray irradiation on the deformation and fracture behavior of human cortical bone. *Bone* **46**, 1475–1485. <https://doi.org/10.1016/j.bone.2010.02.025> (2010).
7. Barth, H. D. *et al.* Characterization of the effects of x-ray irradiation on the hierarchical structure and mechanical properties of human cortical bone. *Biomaterials* **32**, 8892–8904. <https://doi.org/10.1016/j.biomaterials.2011.08.013> (2011).
8. Brent, R. L. The effect of embryonic and fetal exposure to x-ray, microwaves, and ultrasound: Counseling the pregnant and non-pregnant patient about these risks. *Semin. Oncol.* **16**, 347–368 (1989).

9. Faraj, K. & Mohammed, S. Effects of chronic exposure of X-ray on hematological parameters in human blood. *Comp. Clin. Pathol.* **27**, 31–36. <https://doi.org/10.1007/s00580-017-2547-7> (2018).
10. Liang, X., Zhang, J. Y., Cheng, I. K. & Li, J. Y. Effect of high energy X-ray irradiation on the nano-mechanical properties of human enamel and dentine. *Braz. Oral Res.* <https://doi.org/10.1590/1807-3107BOR-2016.vol30.0009> (2016).
11. Haworth, C. D., De Saint-Pern, Y., Clark, D., Trucco, E. & Petillot, Y. R. Detection and tracking of multiple metallic objects in millimetre-wave images. *Int. J. Comput. Vision* **71**, 183–196. <https://doi.org/10.1007/s11263-006-6275-8> (2007).
12. Haworth, C. D., De Saint-Pern, Y., Petillot, Y. R. & Trucco, E. Public security screening for metallic objects with millimetre-wave images. *IET Conference Proceedings*, 1–4 (2005). https://digital-library.theiet.org/content/conferences/https://doi.org/10.1049/ic_20050060.
13. Haworth, C. *et al.* *Image analysis for object detection in millimetre-wave images*. Vol. 5619 ESD (SPIE, 2004).
14. QinetiQ. Security screening visitors at a large London venue. (2018).
15. *Innovations in Defence Support Systems - 2 - Socio-Technical Systems*. 1 edn, (Springer, 2011).
16. Tzydynzhapov, G. *et al.* New real-time sub-terahertz security body scanner. *J. Infrared Millim. Terahertz Waves* **41**, 632–641. <https://doi.org/10.1007/s10762-020-00683-5> (2020).
17. Butavicius, M. A. *et al.* in *Innovations in Defence Support Systems -2: Socio-Technical Systems* (eds Lakhmi C. Jain, Eugene V. Aidman, & Canicious Abeynayake) 183–206 (Springer Berlin Heidelberg, 2011).
18. Thruvision. Profit Protection Product Range.
19. Graham-Rowe, D. Terahertz takes to the stage. *Nat. Photon.* **1**, 75–77. <https://doi.org/10.1038/nphoton.2006.85> (2007).
20. Binstock, J. & Minukas, M. *Developing an Operational and Tactical Methodology for Incorporating Existing Technologies to Produce the Highest Probability of Detecting an Individual Wearing an IED* Masters thesis, Naval Postgraduate School, (2010).
21. Dickson, M. R. *Handheld Infrared camera use for suicide bomb detection: feasibility of use for thermal model comparison* Master of Science thesis, Kansas State University, (2008).
22. Kowalski, M., Kastek, M., Piszczek, M., Życzkowski, M. & Szustakowski, M. *Harmless screening of humans for the detection of concealed objects*. (2015).
23. Kowalski, M., Grudzień, A., Palka, N. & Szustakowski, M. *Face recognition in the thermal infrared domain*. Vol. 10441 ESD (SPIE, 2017).
24. Kowalski, M., Kastek, M. & Szustakowski, M. Concealed objects detection in visible, infrared and terahertz ranges. *Int. J. Comput. Inform. Syst. Control Eng.* **8**, 1632–1627 (2014).
25. Kowalski, M. *et al.* Investigations of concealed objects detection in visible, infrared and terahertz ranges. *Photon. Lett. Pol.* **5**, 167–169. <https://doi.org/10.4302/plp.2013.4.16> (2013).
26. JasimHussein, N., Hu, F. & He, F. Multisensor of thermal and visual images to detect concealed weapon using harmony search image fusion approach. *Pattern Recog. Lett.* **94**, 219–227. <https://doi.org/10.1016/j.patrec.2016.12.011> (2017).
27. Corsi, C. in *Sensors*. (eds Francesco Baldini *et al.*) 37–42 (Springer New York).
28. Kowalski, M. Real-time concealed object detection and recognition in passive imaging at 250GHz. *Appl. Opt.* **58**, 3134–3140. <https://doi.org/10.1364/AO.58.003134> (2019).
29. Slamani, M., Varshney, P. K., Rao, R. M., Alford, M. G. & Ferris, D. in *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348)*. vol. 513, 518–522.
30. Xue, Z., Blum, R. S. & Li, Y. in *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002. (IEEE Cat. No.02EX5997)*. vol. 1192, 1198–1205.
31. Khor, W., Chen, Y. K., Roberts, M. & Ciampa, F. Infrared thermography as a non-invasive scanner for concealed weapon detection. *Defence Secur. Doctoral Symp. 2023* <https://doi.org/10.17862/cranfield.rd.25028030.v2> (2024).
32. Wang, Z., Wan, L., Xiong, N., Zhu, J. & Ciampa, F. Variational level set and fuzzy clustering for enhanced thermal image segmentation and damage assessment. *NDT & E Int.* **118**, 102396. <https://doi.org/10.1016/j.ndteint.2020.102396> (2021).
33. Wang, Z., Zhu, J., Tian, G. & Ciampa, F. Comparative analysis of eddy current pulsed thermography and long pulse thermography for damage detection in metals and composites. *NDT & E Int.* **107**, 102155. <https://doi.org/10.1016/j.ndteint.2019.102155> (2019).
34. Wang, D., Wang, Z., Zhu, J. & Ciampa, F. Enhanced pre-processing of thermal data in long pulse thermography using the Levenberg–Marquardt algorithm. *Infrared Phys. Technol.* **99**, 158–166. <https://doi.org/10.1016/j.infrared.2019.04.009> (2019).
35. Wang, Z., Tian, G., Meo, M. & Ciampa, F. Image processing based quantitative damage evaluation in composites with long pulse thermography. *NDT & E Int.* **99**, 93–104. <https://doi.org/10.1016/j.ndteint.2018.07.004> (2018).
36. Chen, B., Wang, W. & Qin, Q. Robust multi-stage approach for the detection of moving target from infrared imagery. *Opt. Eng.* **51**, 067006 (2012).
37. Cotič, P., Jagličić, Z., Niederleithinger, E., Stoppel, M. & Bosiljkov, V. Image fusion for improved detection of near-surface defects in NDT-CE using unsupervised clustering methods. *J. Nondestruct. Eval.* **33**, 384–397. <https://doi.org/10.1007/s10921-014-0232-1> (2014).
38. Hong, J., Cheng, H., Zhang, Y.-D. & Liu, J. Detecting cerebral microbleeds with transfer learning. *Mach. Vision Appl.* **30**, 1123–1133. <https://doi.org/10.1007/s00138-019-01029-5> (2019).
39. Muazzam, M. *et al.* Transfer learning assisted classification and detection of Alzheimer’s disease stages using 3D MRI scans. *Sensors* <https://doi.org/10.3390/s19112645> (2019).
40. Liang, G. & Zheng, L. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Comput. Methods Progr. Biomed.* **187**, 104964. <https://doi.org/10.1016/j.cmpb.2019.06.023> (2020).
41. Shin, H. C. *et al.* Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **35**, 1285–1298. <https://doi.org/10.1109/TMI.2016.2528162> (2016).
42. Gezimati, M. & Singh, G. in *2022 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*. 1–6.
43. Gao, Y. & Mosalam, K. M. Deep transfer learning for image-based structural damage recognition. *Comput. Aided Civil Infrastruct. Eng.* **33**, 748–768. <https://doi.org/10.1111/mice.12363> (2018).
44. Dawei, W. *et al.* Recognition pest by image-based transfer learning. *J. Sci. Food Agric.* **99**, 4524–4531. <https://doi.org/10.1002/jsfa.9689> (2019).
45. Mukti, I. Z. & Biswas, D. in *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*. 1–6.
46. Shabbir, A. *et al.* Satellite and scene image classification based on transfer learning and fine tuning of ResNet50. *Math. Probl. Eng.* **2021**, 5843816. <https://doi.org/10.1155/2021/5843816> (2021).
47. Huang, Y., Jiang, Z., Lan, R., Zhang, S. & Pi, K. Infrared image super-resolution via transfer learning and PSRGAN. *IEEE Signal Process. Lett.* **28**, 982–986. <https://doi.org/10.1109/LSP.2021.3077801> (2021).
48. Hu, J., Zhao, Y. & Zhang, X. in *2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC)*. 1–4.
49. Arthur, D. & Vassilvitskii, S. in *SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms* 1027–1035 (2007).
50. Bezdek, J. C. *Pattern Recognition with Fuzzy Objective Function Algorithms* (Springer, 1981).
51. Panda, S., Sahu, S., Jena, P. & Chattopadhyay, S. in *Advances in Computer Science, Engineering & Applications*. (eds David C. Wyld, Jan, Z., & Dhinakaran, N) 451–460 (Springer).
52. Cebeci, Z. & Yildiz, F. Comparison of K-means and fuzzy C-means algorithms on different cluster structures. *J. Agric. Inform.* <https://doi.org/10.17700/jai.2015.6.3.196> (2015).

53. The MathWorks Inc. MATLAB version: 9.12.0 (R2022a). (2022).
54. He, K., Zhang, X., Ren, S. & Sun, J. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778.
55. Barry-Straume, J., Tschannen, A., Engels, D. W. & Fine, E. An evaluation of training size impact on validation accuracy for optimized convolutional neural networks. *SMU Data Sci. Rev.* **1**(4), 12 (2018).
56. Cho, J., Lee, K., Shin, E., Choy, G. & Do, S. How much data is needed to train a medical image deep learning system to achieve necessary high accuracy?. *ICLR* <https://doi.org/10.48550/arXiv.1511.06348> (2016).
57. Deng, J. *et al.* in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 248–255.
58. Yamashita, R., Nishio, M., Do, R. K. G. & Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **9**, 611–629. <https://doi.org/10.1007/s13244-018-0639-9> (2018).
59. Mandrekar, J. N. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* **5**, 1315–1316. <https://doi.org/10.1097/JTO.0b013e3181ec173d> (2010).
60. Zhang, H., Hu, T. L. & Zhang, J. C. Surface emissivity of fabric in the 8–14 μm waveband. *J. Text. Inst.* **100**, 90–94. <https://doi.org/10.1080/00405000701692486> (2009).

Acknowledgements

The authors acknowledge funds provided by the UK Defence and Security Accelerator (DASA) [Grant Number ACC2022360] for the IR SCREEN project.

Author contributions

Conceptualization, W.K., Y.K.C., M.R., F.C.; methodology, W.K., Y.K.C.; formal analysis, W.K.; investigation, W.K.; resources, F.C.; data curation, W.K.; writing—original draft preparation, W.K.; writing—review and editing, Y.C.K., M.R., F.C.; visualization, W.K.; supervision, M.R. and F.C.; project administration, F.C.; funding acquisition, F.C. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to F.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024