Časopis Pomorskog fakulteta Kotor – Journal of Maritime Sciences (JMS)
god. 24, br. 1/2023; Vol. 24, No. 1/2023; May, 2023

DOI: https://doi.org/10.56080/jms230504 UDC: 339.543:629.5.08
Original scientific paper

Automatic Detection of Visual Changes between X-Ray Images to Support Antifraud Customs Control at Ports – A Preliminary Study*

Andrea Chezzi, Mattia Colucci, Radmila Gagic, Danilo Martino, Claudio Pascarelli†, Adriana Pettinicchio

Abstract: One of the goals of customs authorities is to identify, at borders, cargos that do not match their declaration, contain illegal items, or pose a hazard to society. Customs X-ray inspection procedures enable the detection of suspicious cargos and are an excellent support tool for customs officials. The ISACC project, funded under the Interreg IPA Cross-border Cooperation Italy-Albania-Montenegro Programme aims at developing a web platform that integrates data, coming from heterogeneous technologies and systems, in order to provide a rich information base supporting customs authorities during antifraud controls. In this paper, we propose a preliminary study, based on the SIFT algorithm, for the automatic detection of visual changes between scanner X-Ray images, that are part of this information base.

Keywords: Customs, Antifraud, X-Ray, Computer vision, SIFT.

1. Introduction

The current age of globalization with its advances in transportation and information technology has increased trade across the world and the freedom in which these trades are carried out [1]. The total gross weight of goods handled in EU ports in 2021 was estimated at 3.5 billion tons, a 4% increase compared with 2020 [2]. The expansion of global transportation raises security concerns because any container or truck could be used by malicious actors to smuggle restricted or prohibited items across borders. The methods employed by criminals include concealment of undeclared goods amongst a legitimate cargo or in the fabric of the container itself (e.g. floor, refrigeration unit) [3].

* An earlier version of this paper was presented at the 2nd Kotor International Maritime Conference – KIMC 2022, Kotor, Montenegro.

† Corresponding author

One of the goals of customs authorities is to identify cargos that don't match their declaration, contain illegal items, or pose a hazard to society. The smuggling of fake and pirated goods hurts a country's economy since taxes are not paid that would otherwise be used to benefit society. The existence of counterfeit products is a significant crime problem in the twenty-first century [4]. Focusing on Italy, drugs and cigarettes represent the categories of goods that are most smuggled [5]. As for unexpected, illicit and harmful products, "light" weapons, trafficking in natural resources and the illegal trade in wildlife are the most impacting plagues in the EU beside narcotics [6].

Technological advances and smart policies are required to facilitate the inspection and achieve integrated security [1]. Customs X-ray inspection procedures enable the detection of suspicious cargos and are an excellent support tool for customs officials. Usually the scanner images are stored in a central reference database, which contains X-ray images of legal and illegal cargos in a manufacturer-independent format. The data in the reference database can be shared with other customs administrations in order to facilitate the exchange of information by, for example, comparing cargos in transit from one inspection point to another.

The ISACC project, funded under the Interreg IPA Cross-border Cooperation Italy-Albania-Montenegro Programme, fits within this context, since it aims to develop a web platform that integrates data coming from heterogeneous technologies and systems, already available to the customs authorities of three countries (Italy, Albania, Montenegro), in order to provide a rich information base called Custom Footprint (CF) supporting customs authorities during antifraud controls. The CF is a set of data regarding a specific target, such as a container with related goods, that is created at the first customs control point (e.g. at the time of export) and is tracked till the final destination (e.g. at the time of import inspections). Customs officers of three selected pilot sites (i.e. Port of Bari in Italy, Port of Durres in Albania, Port of Bar in Montenegro) can check, by means of the platform, the invariance of the information of the CF in each intermediate stage defined as checkpoint (e.g. customs control point at the three ports). One, and probably the most important information, within a CF, is represented by scanner images. The ISACC platform allows customs officers to compare two scanner images in order to find differences and generate an alert when needed. In this paper, we propose a preliminary study based on the SIFT algorithm (Scale-Invariant Feature Transform) for the automatic detection of visual changes between X-Ray images in order to support antifraud customs control.

In the next section related work on this topic will be discussed, in order to evaluate what scientific advancement has achieved so far and to identify specific areas where further improvements are required. In section 3, the methodology applied in this study is described and, in section 4, the preliminary obtained results are discussed. Finally, a conclusions section ends the paper.

2. Theoretical background and related works

Computer vision fields are rising in the recent past. These techniques include the image matching that plays an important role in many applications. In the evolution of image matching techniques, a lot of algorithms have been proposed in the literature [7-9].

The discipline of computer vision addresses the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, multidimensional data from a 3D scanner, medical scanning devices, or - as it happens in this case - from X-ray scanners. The technological discipline of computer vision seeks to apply its theories and models to the construction of computer vision systems. Such autonomous systems could perform some of the tasks which the human visual system can perform, and even surpass it in many cases.

Normally, the use of computer vision techniques involves a preliminary step of image acquisition, where images and large sets can be acquired in real-time through video, photos or 3D technology for analysis. Once visual information is acquired, the step of image processing kicks in, in which either machine learning models or conventional algorithms are used to automate much of this process. However, in the case of machine learning, the models are often trained by first being fed a consistent number of labelled or pre-identified images. The final step is the interpretative step, where an object is identified or classified.

Among the many possible applications, image recognition and matching algorithms have also been applied to analyse X-ray images of cargo scanning. Most studies focused on recognizing a specific target object in a cargo.

Jaccard et al. [10] used, for example, Deep Learning to detect concealed cars in X-ray cargo images: they proposed an algorithm based on trained-from-scratch Convolutional Neural Networks. In [3] the authors, on the other hand, implemented a framework to identify an empty cargo, by handling the task as a binary classification problem.

Another field of application relates to the identification of illicit or banned items. Visser et al. [1], for example, defined an automated target

recognition function to analyse scanner X-ray images after: they defined an algorithm to detect certain types of goods such as cigarettes, weapons and drugs in the freight of a container.

A Deep Learning approach for the detection of small threats in X-ray cargo images was presented by Jaccard et al. [11], in this study authors have defined three algorithms based on CNN and Random Forest approaches and demonstrate the major efficiency of the algorithms based on a 19-layer CNN.

In [12] the authors propose an object detection method for efficiently detecting contraband items in both cargo and baggage for X-ray security scans. The proposed network, MFA-net, consists of three plug-and-play modules, including the multiscale dilated convolutional module, fusion feature pyramid network, and auxiliary point detection head. Authors tested the performance of the MFA-net on two large-scale X-ray security image datasets from different domains: a Security Inspection X-ray (SIXray) dataset in the baggage domain and a cargo dataset (CargoX).

Ahmed et al. propose in [13] some algorithms for automatic historical comparison of scanned vehicles. The presented system uses a database of scanned vehicles. Each time a new vehicle is scanned, its license plate number is extracted using a license plate reader, and that number is used as the primary key to retrieve matching vehicles from the database. The two vehicles are then segmented using a model-based segmentation approach and certain points of interest are identified. Image registration is performed on the two images to align them. Intensity profiles for the two images are also normalized, and then the two images are compared to find differences. False alarms are then removed using knowledge-based rules. The system has been tested on images acquired using a deployed scanner and produced satisfactory results.

Chen et al. in their work [14] discuss the importance of improving security inspection capabilities in public transportation due to the impact of emerging terrorist attacks. With the rapid development of X-ray detection technology, different X-ray inspection techniques have been researched and developed to detect hidden explosives and contrabands. The article provides a general review of these techniques, including X-ray transmission imaging, backscatter imaging, phase-contrast imaging, spectral imaging, CT reconstruction, and X-ray diffraction. Each technique is explained with its basic physical mechanisms, research progress, and application features. The article also discusses new technologies and applications that show great potential for security inspections.

Unlike what has been previously discussed in most of the literature, the present study does not focus on finding specific target items/goods in a

cargo but rather on quantifying the percentage difference between two X-ray images of the same cargo, as a decision support for the customs operator who can then decide to proceed with a more thorough inspection of the cargo or not. This work presents similar goals as what is discussed in [15], but with different means to achieve the results.

Also, the current paper focuses on a specific aspect of computer vision where the use of deep learning models is not strictly required or even recommended, since traditional non-ML techniques, such as SIFT, not only can handle this task in a more computationally efficient way but they do not inherently require a large dataset of image data as input in order to work.

3. Method

3.1. Dataset

In order to preliminary validate the algorithms for X-ray images pre-processing and comparison, and the relevant code that has been developed, a synthetic dataset of images has been created. To create it the team started from four real cargo X-ray images with a resolution ranging from 716 x 402 pixels (total area: 287.832 pixels) to 1379 x 397 pixels (total area: 547.463 pixels). For each of these images, ten other images that had differences (in varying percentages) when compared to the original source image, were manually created through image editing software. Five of these images were obtained by subtraction of material in the truck trailer, five others by substitution of the material.

Figure 1 shows one of the original four images and the ten images that were generated from it.

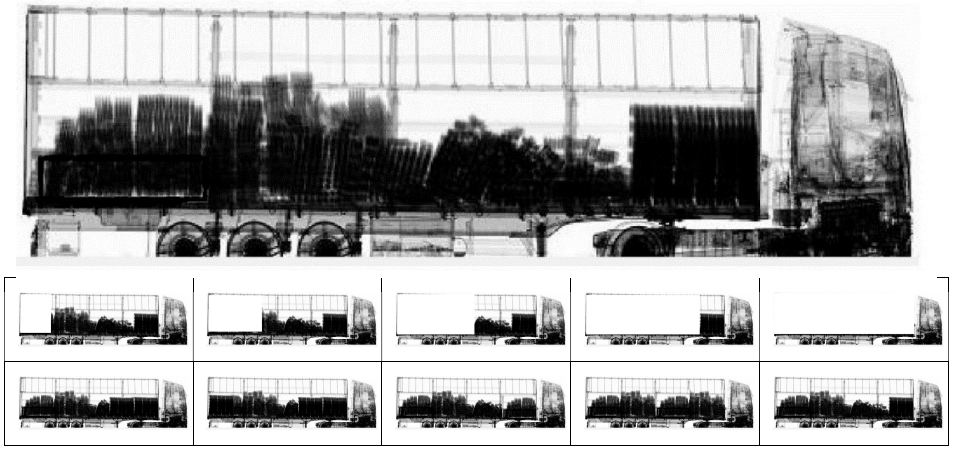

Fig. 1 – *Example of X-ray scanner images and its ten modified variants*

For the 40 images thus generated, the modified area was annotated, in absolute value (pixels) and percentage, as shown in Tables 1-4. The maximum variance in the whole sample of 40 images is 56.5 percent of the image area, which corresponds to the entire volume of transported goods, the average variance is 14.1 percent, and the minimum variance is 0.3 percent.

3.2. Pre-processing

The pre-processing phase is of paramount importance to best perform comparative images analysis: it allows to fit an image over an image target in order to improve the image comparison result.

The pre-processing phase includes the automatic alignment of the two images that need to be compared using the well-known SIFT algorithm (Scale-Invariant Feature Transform) [15]. Scale and rotation invariance are the SIFT best characteristic; scale invariance is guaranteed using DoG (Difference of Gaussian) function.

The SIFT algorithm can be subdivided in three main steps: 1) keypoint detection, 2) descriptor establishing, and 3) image feature matching. In the first phase, SIFT uses grayscale information of an image to identify the image keypoints. In the second phase SIFT uses local information to describe each keypoint. In the last phase SIFT uses a descriptor for image feature matching [16].

The implemented SIFT based algorithm takes as input the couple of images that need to be compared in order to find the differences, one of them is used as reference of the alignment process and the other one is geometrically transformed in order to match as much as possible the reference image. This is done to get the maximum homogeneity possible between the two images during the comparative analysis.

In Figure 2 the image (a) is the image to be aligned and scaled according to the reference (b) image. In the (c) image SIFT keypoints are plotted and the (d) image is the pre-processing result: the image to be compared with (b).

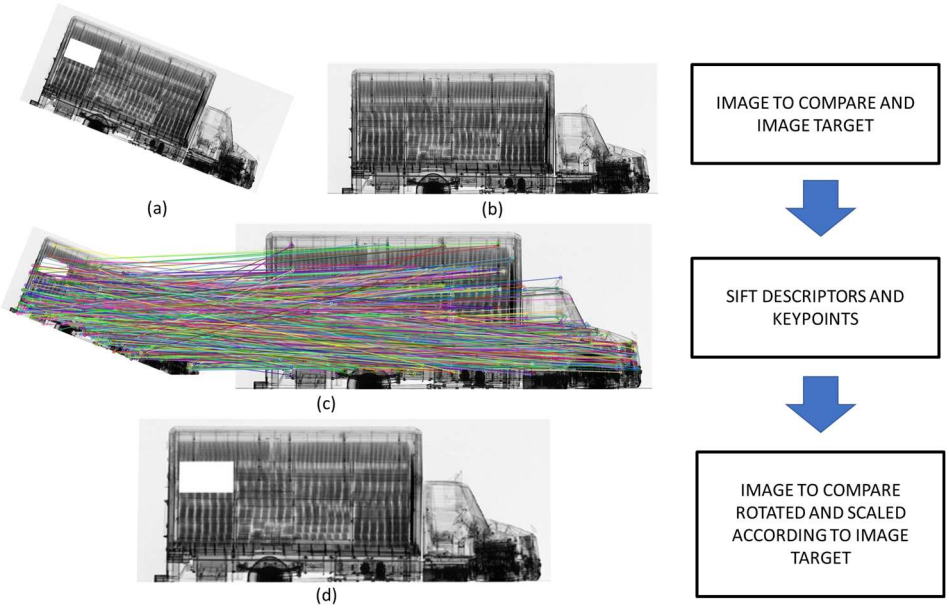

Fig. 2 - *X-ray scanner images pre-processing*

3.3. Image comparison

In this section the comparative images analysis is described. After the pre-processing phase the couple of images has the same scale and rotation, so it's possible to actually perform their comparison.

The Feature Matching algorithms, previously mentioned for pre-processing phase, can also be used to find as many common graphic features as possible between the images, and use the spatial sparsity of the found common descriptors as an indicator of the presence (or absence) of differences between the two images.

The high density of descriptors found in a certain area indicates absence of differences and hence homogeneity. On the other hand, low density or absence of descriptors indicates a high probability of differences which may lie in that particular region.

The image comparison algorithm is based on the SIFT algorithm.

The algorithm can be divided in three steps, as reported in the Figure 3:

1. keypoints and descriptors identification;
2. descriptors matching;
3. similarity percentage calculation.

In the first phase the SIFT algorithm is used to identify features for each image.

The SIFT descriptor represents the input for the second phase that allows to identify similar features between the target image and the image to be compared. In this phase a Brute-force (BF) descriptor matcher is used; the BF matcher iterates over the first image descriptor and keypoints and, for each descriptor in the first set, finds the closest descriptor in the second set by trying each one. The BF matcher is associated with the Cross Checking validator to ensure a consistent feature matching, so the matcher returns only those matches with value (*i,j*) such that *i*-th descriptor in the image target has *j*-th descriptor in the image to be compared as the best match and vice-versa.

The last step of the algorithm is the similarity percentage calculator: this formula is based on the number of features detected by the BF matcher and the images keypoints.

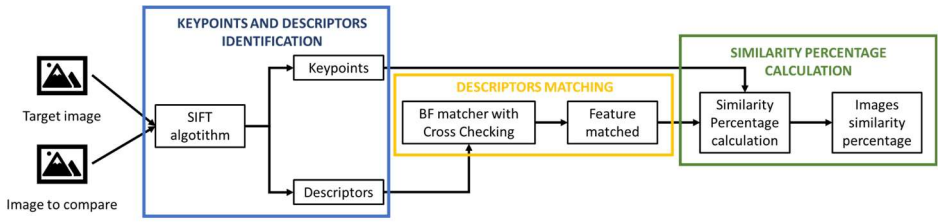

Fig. 3 – *Images comparison workflow*

4. Results

The performance of the Feature Matching algorithms for image comparison was estimated through two separate error indexes: Mean Squared Error (MSE) and Root Mean Square Error (RMSE). These error measures are the most popular in various domains [17, 18]

Mean squared error (MSE) measures the amount of error in statistical models. It assesses the average squared error (*e*), i.e the difference between the observed and predicted values, i.e the real percentage of difference between the compared pairs of images and the percentage of difference estimated through the proposed method based on the Feature Matching algorithm. When a model has no error, the MSE equals zero. As model error increases, its value increases.

The formula for MSE is the following.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \quad (1)$$

Alternatively, RMSE, that has been used as a standard statistical metric to measure model performance in several research fields [19], is calculated easily by taking the square root of MSE.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \quad (2)$$

As for MSE, the RMSE has frequently been used as an evaluation indicator to assess the reliability and accuracy of the estimated parameters hence summarizing the overall error of a model [20].

In the following Tables 1-4 the individual errors for each of the 40 image comparisons are reported.

Table 1 – *First images sub-set and relevant data*

Image	Total area in pixels	Modified area in pixels	Modified area (M)	Estimated difference (E)	Error (E-M)^2
010 (original)	521.180				
010a	521.180	60.828	11,67%	1,67%	1,000%
010b	521.180	98.688	18,94%	4,56%	2,067%
010c	521.180	159.030	30,51%	4,50%	6,767%
010d	521.180	230.175	44,16%	8,74%	12,549%
010e	521.180	294.640	56,53%	5,96%	25,577%
010f	521.180	28.480	5,46%	4,76%	0,005%
010g	521.180	56.960	10,93%	9,15%	0,032%
010h	521.180	41.236	7,91%	6,36%	0,024%
010i	521.180	105.248	20,19%	16,44%	0,141%
010j	521.180	1.776	0,34%	0,00%	0,001%

Table 2 – *Second images sub-set and relevant data*

Image	Total area in pixels	Modified area in pixels	Modified area (M)	Estimated difference (E)	Error (E-M)^2
020 (original)	287.832				
020a	287.832	2.262	0,79%	0,43%	0,001%
020b	287.832	17.150	5,96%	2,45%	0,123%
020c	287.832	21.534	7,48%	4,23%	0,106%
020d	287.832	30.821	10,71%	5,33%	0,289%
020e	287.832	58.308	20,26%	10,21%	1,010%
020f	287.832	10.089	3,51%	11,72%	0,675%
020g	287.832	16.416	5,70%	19,61%	1,934%
020h	287.832	23.664	8,22%	17,63%	0,885%
020i	287.832	1.443	0,50%	0,56%	0,000%
020j	287.832	756	0,26%	1,61%	0,018%

Table 3 – *Third images sub-set and relevant data*

Image	Total area in pixels	Modified area in pixels	Modified area (M)	Estimated difference (E)	Error (E-M)^2
030 (original)	547.463				
030a	547.463	8.858	1,62%	1,57%	0,000%
030b	547.463	36.494	6,67%	2,90%	0,142%
030c	547.463	71.960	13,14%	5,87%	0,529%
030d	547.463	134.919	24,64%	6,54%	3,278%
030e	547.463	261.320	47,73%	16,90%	9,507%
030f	547.463	22.794	4,16%	6,67%	0,063%
030g	547.463	34.191	6,25%	1,92%	0,187%
030h	547.463	89.001	16,26%	22,60%	0,402%
030i	547.463	122.670	22,41%	24,30%	0,036%
030j	547.463	1.980	0,36%	0,71%	0,001%

Table 4 – *Fourth images sub-set and relevant data*

Image	Total area in pixels	Modified area in pixels	Modified area (M)	Estimated difference (E)	Error (E-M)^2
040 (original)	378.099				
040a	378.099	9.271	2,45%	1,76%	0,005%
040b	378.099	25.870	6,84%	3,20%	0,133%
040c	378.099	72.280	19,12%	4,00%	2,285%
040d	378.099	118.428	31,32%	6,38%	6,221%
040e	378.099	171.698	45,41%	7,94%	14,041%
040f	378.099	9.975	2,64%	5,00%	0,056%
040g	378.099	13.674	3,62%	5,92%	0,053%
040h	378.099	26.069	6,89%	11,18%	0,184%
040i	378.099	51.975	13,75%	24,43%	1,141%
040j	378.099	77.562	20,51%	31,94%	1,306%

The resulting values of MSE and RMSE are:

$$MSE = \frac{1}{40}\sum_{i=1}^{40} e_i^2 = 2{,}32\ \% \tag{3}$$

$$RMSE = \sqrt{\frac{1}{40}\sum_{i=1}^{40} e_i^2} = 15{,}23\ \% \tag{4}$$

If look at the correlation between images' modified area and the square error (Figure 4), it can be seen that the Feature Matching algorithm performs better if the difference between the original and the modified image is lower than 20 %. In this case the MSE is 0.43 % that becomes even lower if we only consider the couple of images having differences lower than 10 %: in this case the MSE is 0.22 %.

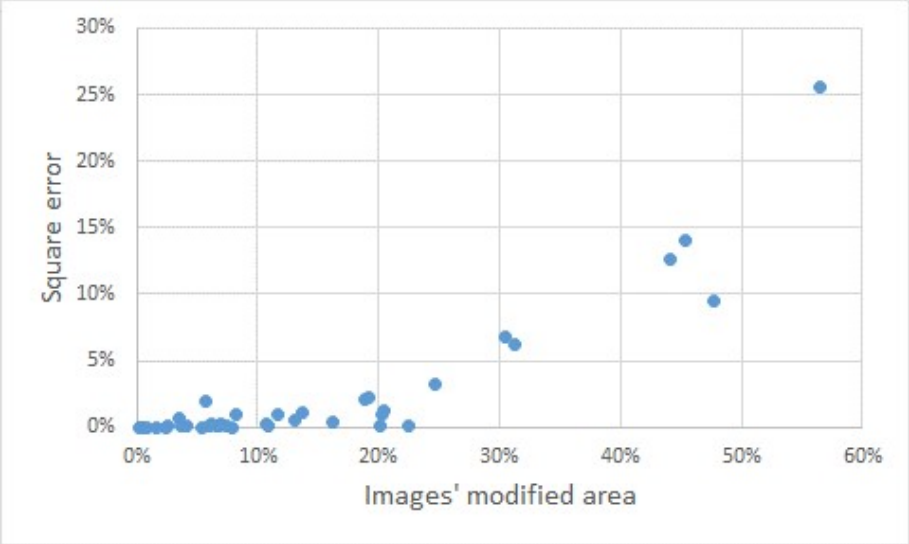

Fig. 4 – *Correlation between images' modified area and the measured square error*

5. Conclusions

Every year the global international trade in counterfeit, pirated and illegal goods amounts to hundreds of billion euros. Customs authorities working at borders play a vital role in contrasting this phenomenon especially in those that have presented problems in this regard for decades, such as those between Italy, Albania and Montenegro.

Within the ISACC project an IT platform, supporting antifraud customs controls, has been designed and developed to integrate data coming from heterogeneous technologies and systems in order to provide a rich information base that we called Custom Footprint (CF). Customs authorities can use this CF and search for its variances, that would be a possible sign of a fraud, at intermediate customs control points, i.e. the ports of Bari, Durres and Bar if we refer to the ISACC project. In order to compare scanner X-ray images that are part of the CF, in this paper it is proposed a possible approach, based on the SIFT algorithm (Scale-Invariant Feature Transform) and some preliminary results obtained with a synthetic dataset.

As shown in the relevant section, the results can be considered encouraging, especially where the differences between two images is below the 20 % threshold. This represents the case in which the proposed method is most useful: substantial differences between two images can in fact be

easily detected by sight by a customs operator without the support of a computer tool.

Future work will be directed toward further improving the developed method and testing it on a much larger number of samples (both synthetic and real). One possible solution that will be implemented and tested will be to divide the images into smaller sections (tiles) so as to have not only an indicator of the difference between the two images but also to identify the region where this difference exists. The adoption of another "matcher" still based on SIFT, but adopting a method for skimming the pairs of features found will also be evaluated.

6. Acknowledgments

The study was carried out within the “Innovative Systems to enhance Antifraud Customs Controls'' (ISACC) project funded under the Interreg IPA CBC Italy-Albania-Montenegro 2014–2020 programme (project ID P.A. 4 / S.O. 4.1 - N. 365). The authors gratefully acknowledge the Italian, Albanian and Montenegrin custom administrations for their support during the project.

References

[1] Visser, W., Schwaninger, A., Hardmeier, D., Flisch, A., Costin, M., Vienne, C., Sukowski, F., Hassler, U., Dorion, I., Marciano, A., Koomen, G., Slegt, M., & Canonica, A. (2016). Automated comparison of X-ray images for cargo scanning. In 2016 IEEE International Carnahan Conference on Security Technology (ICCST) (pp. 1-8). IEEE.

[2] https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20221116-3 (accessed 15 December 2022)

[3] Jaccard, N., Rogers, T. W., Morton, E. J., & Griffin, L. D. (2016). Tackling the X-ray cargo inspection challenge using machine learning. In Anomaly Detection and Imaging with X-Rays (ADIX) (9847, 131-143). SPIE.

[4] Taraba, P., Hoke, E., & Marada, J. (2020). Means of transporting counterfeits to the European Union. Chemical Engineering Transactions.

[5] Talarico, L., & Zamparini, L. (2017). Intermodal transport and international flows of illicit substances: Geographical analysis of smuggled goods in Italy. Journal of transport geography, 60, 1-10.

[6] Nanfosso, R. T., & Hadjitchoneva, J. (2021). The European Union Facing the Challenges of Globalisation. Economic Studies, 30(5).

[7] Jayanthi, N., & Indu, S. (2016). Comparison of image matching techniques. International journal of latest trends in engineering and technology, 7(3), 396-401.
[8] Heipke, C. (1996). Overview of image matching techniques. In OEEPE Workshop on the Application of Digital Photogrammetric Workstations.
[9] Zitova, B., & Flusser, J. (2003). Image registration methods: a survey. Image and vision computing, 21(11), 977-1000.
[10] Jaccard, N., Rogers, T. W., Morton, E. J., & Griffin, L. D. (2017). Detection of concealed cars in complex cargo X-ray imagery using deep learning. Journal of X-ray Science and Technology, 25(3), 323-339.
[11] Jaccard, N., Rogers, T. W., Morton, E. J., & Griffin, L. D. (2015). Using deep learning on X-ray images to detect threats. In Proceedings Cranfield Defence and Security Doctoral Symposium (pp. 1-12).
[12] Viriyasaranon, T., Chae, S. H., & Choi, J. H. (2022). MFA-net: Object detection for complex X-ray cargo and baggage security imagery. Plos one, 17(9), e0272961.
[13] Ahmed, W. M., Zhang, M., & Al-Kofahi, O. (2011). Historical comparison of vehicles using scanned x-ray images. In 2011 IEEE Workshop on Applications of Computer Vision (WACV) (pp. 288-293). IEEE.
[14] Chen, Z. Q., Zhang, L., & Jin, X. (2017). Recent progress on X-ray security inspection technologies. Chinese Science Bulletin, 62(13), 1350-1365.
[15] Burger, W., & Burge, M. J. (2022). Scale-invariant feature transform (SIFT). In Digital Image Processing (pp. 709-763). Springer, Cham.
[16] Wu, J., Cui, Z., Sheng, V. S., Zhao, P., Su, D., & Gong, S. (2013). A Comparative Study of SIFT and its Variants. Measurement science review, 13(3).
[17] De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. International journal of forecasting, 22(3), 443-473.
[18] Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International journal of forecasting, 22(4), 679-688.
[19] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). Geoscientific Model Development Discussions, 7(1), 1525-1534.
[20] Wang, J., Zhao, J., Liu, Z., & Kang, Z. (2021). Location and estimation of multiple outliers in weighted total least squares. Measurement, 181, 109591.

Submitted: 29/12/2022
Accepted: 27/03/2023

Andrea Chezzi
University of Salento,
7 Piazza Tancredi, 73100 Lecce (IT)
andrea.chezzi@unisalento.it

Mattia Colucci
University of Salento,
7 Piazza Tancredi, 73100 Lecce (IT)
mattia.colucci@unisalento.it

Radmila Gagic
University of Montenegro,
2 Cetinjski Put, 81000 Podgorica (ME)
radmilag@ucg.ac.me

Danilo Martino
Advantech s.r.l.,
Via per Monteroni, 73100 Lecce (IT)
danilo.martino@advantechsrl.com

Claudio Pascarelli
University of Salento,
7 Piazza Tancredi, 73100 Lecce (IT)
claudio.pascarelli@unisalento.it

Adriana Pettinicchio
Freelance Software Engineer
adriana.pettinicchio@gmail.com