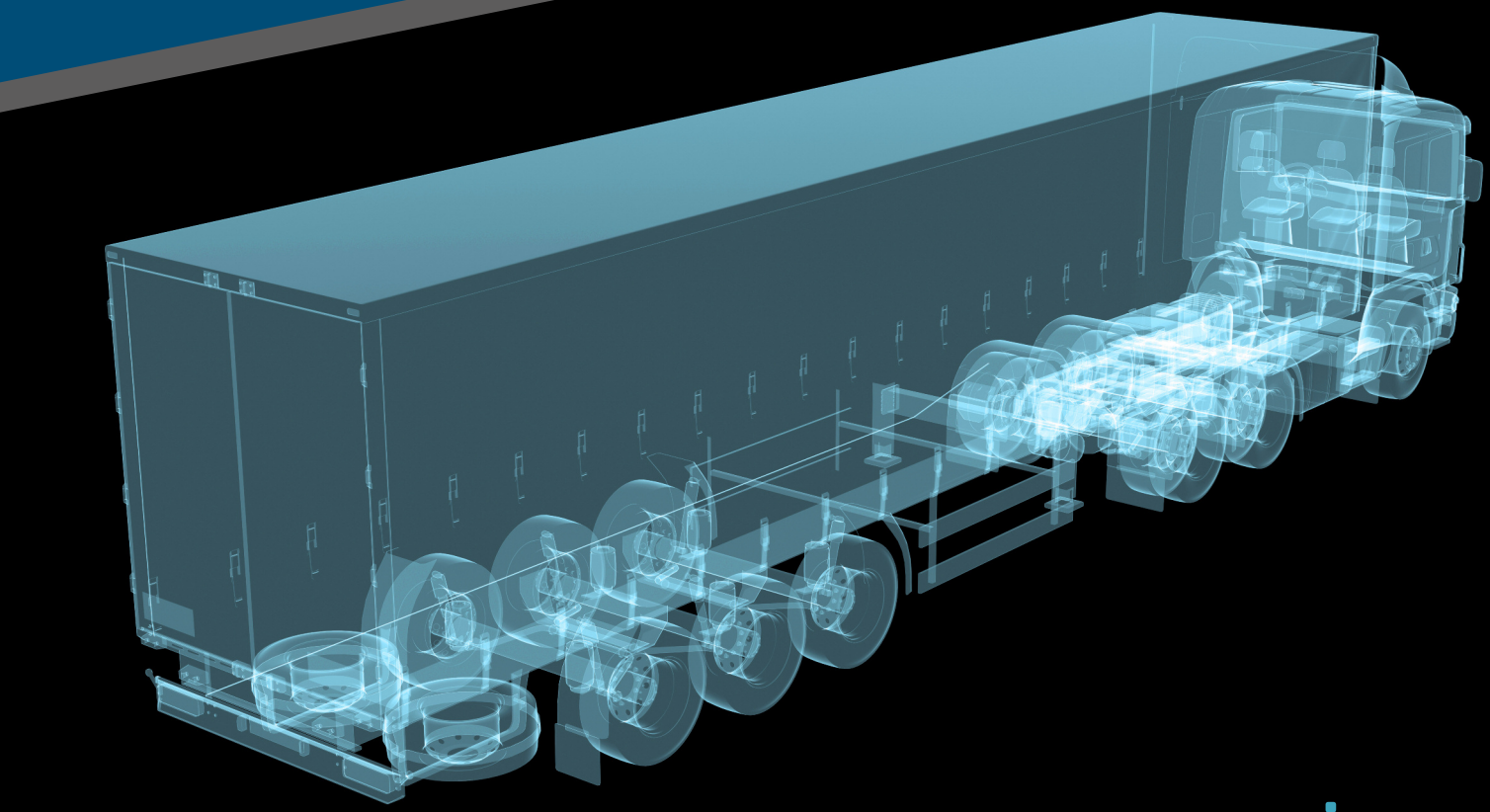
ASSISTED TARGET RECOGNITION FOR X-RAY CARGO SCANNING



www.acxis.eu

Empa
Materials Science and Technology

list
CEA tech

Fraunhofer
EZRT

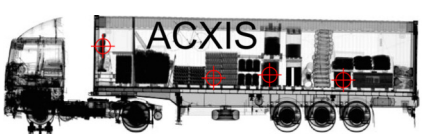
Customs Administration of the Netherlands
Tax and Customs Administration

smiths detection
bringing technology to life

CASRA

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössische Zollverwaltung EZV
Administration fédérale des douanes AFD
Amministrazione federale delle dogane AFD



Validation of improved customs detection procedures

How useful are Assisted Target Recognition algorithms and training for improving the detection performance of customs officers?

Authors: Yanik Sterchi, Wicher Visser

Automated Comparison of X-ray Images for Cargo Scanning (ACXIS) is a research project funded by the European Union under the 7th Framework Program. The ACXIS consortium (including CASRA) conducts research and development for improving effectiveness and efficiency of cargo screening. As part of this, ACXIS partners are developing Assisted Target Recognition (ATR) methods to automatically detect potentially illicit goods[1] by means of dedicated algorithms. For this purpose, first the Dutch Customs Organization (DTCA) and the Swiss Federal Customs Administration (FCA) defined several illicit goods detection scenarios. The consortium then prioritized and selected scenarios, resulting in four ATRs being developed with Smiths Detection in the lead. As an example, methods for cigarette detection based on texture analysis have been developed and show excellent first results. While the overall project has been described in the CASRA Newsletter issue 12 (March 2016), the present article is focusing on the validation study of the ACXIS project.

How effective ATRs are does not only depend on the performance of the automatic algorithms. It is also important, how the information provided by the ATRs is implemented into the screening process. The idea within ACXIS was that ATRs support visual inspection of customs officers by directing their attention on suspicious areas in the X-ray image and help decision making (red frame in Figure 1). In this implementation scenario, the screening officer still makes the final call on whether the cargo needs further inspection. Therefore, whether and how much the detection performance of the screener improves due to ATRs, depends on a variety of factors. Research on human–machine interaction in various professions revealed that the benefit of automated support systems is influenced by the officers' experience with the system ([1]), their expertise in the task ([2]), and the occurrence of hits and false alarms[2] generated by the system (e.g. [1], [3]). Research has also shown that automated support systems do not always lead to better performance (e.g. [4]). A key component of the ACXIS project was therefore to evaluate the designed implementation of the ATRs. Since the officers' experience with the system and their expertise in the task can influence the benefit of an automated system, the successful use of the ATRs might depend on whether screeners are trained and had the chance to gain experience with the ATRs. Computer-based training has repeatedly been shown to increase detection performance in X-ray image inspection of air passenger bags (e.g. [5], [6]) and one study also showed its benefit in cargo screening ([7]). The ACXIS validation study therefore also aimed to investigate the influence of training. The validation study was designed with following goals:

- Replication of the previously found effects of computer-based training on detection performance in cargo screening
- Evaluate whether ATRs lead to an increase in detection performance
- Investigate whether the benefit of the ATRs is influenced by the computer-based training

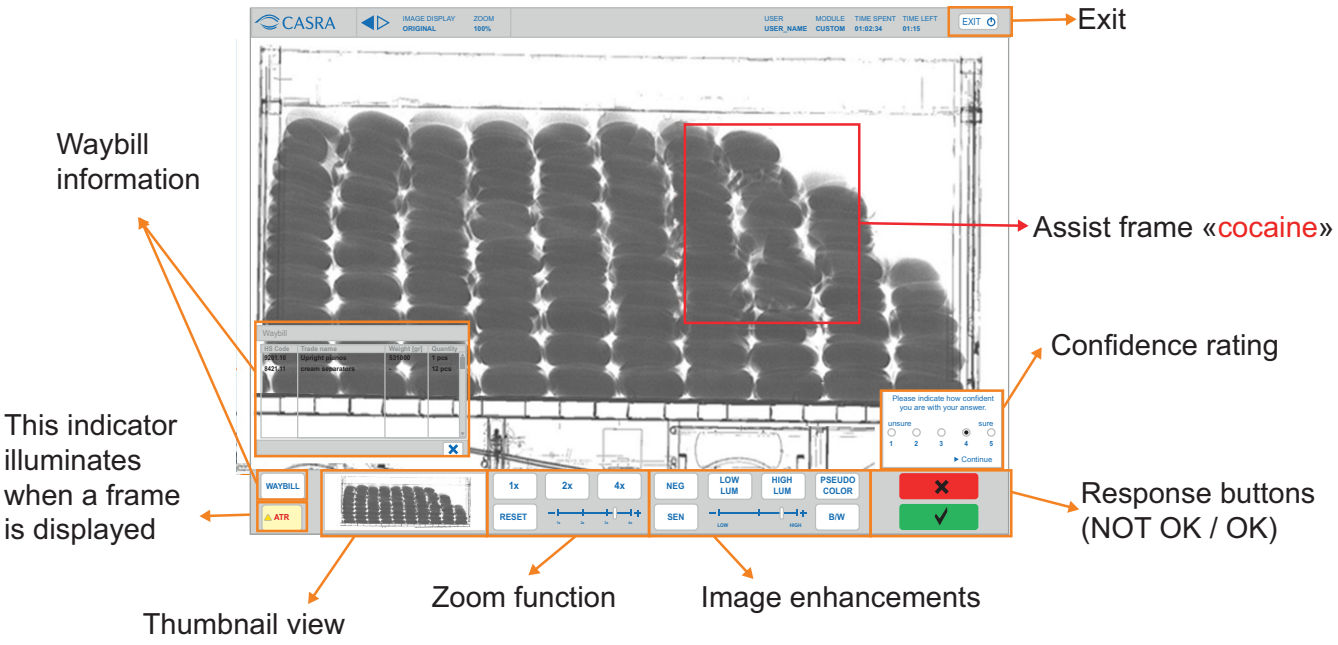


Figure 1: Simulator interface for training and testing with ATRs

METHOD: SIMULATOR

For the validation study, CASRA has developed a new platform for cargo X-ray image interpretation training and testing, called the Customs X-Ray Simulator (described in the CASRA Newsletter issue 13, July 2016).

ATRS

In the validation study, four assist functions have been evaluated. ATRs 1-3 aimed at detecting specific potentially illicit goods within the freight: hidden cigarettes (ATR1), weapons (ATR2) and narcotics (ATR3). A fourth ATR detected anomalies in the container structure (ATR4), and was therefore able to detect any sort of potentially illicit goods, but only in the container structure and not within the freight (Figure 2). For a better control over the experimental conditions, the officers were trained and tested separately for the location-specific ATR4 and the target-specific ATRs 1-3. The detection algorithms of the ATRs have been still under development when the validation study was conducted, and the alarms of the ATRs were therefore set manually by screening experts. All four ATRs were designed to detect 75% of their respective targets, while 25% of the X-ray images produced a false alarm.

CONTENT

To create the training and testing material, the Dutch Customs Organization (DTCA) provided images of cargo recorded with a HCV scanner (also appropriate waybills were included but no top views of the X-ray scans). Empa and DTCA recorded cigarettes, weapons, and mockup narcotics, which were merged into cargo images by CASRA, using a merging tool from CEA. Officers from DTCA and Swiss Federal Customs Administration (FCA) reviewed the images for quality (these officers were excluded from the study).

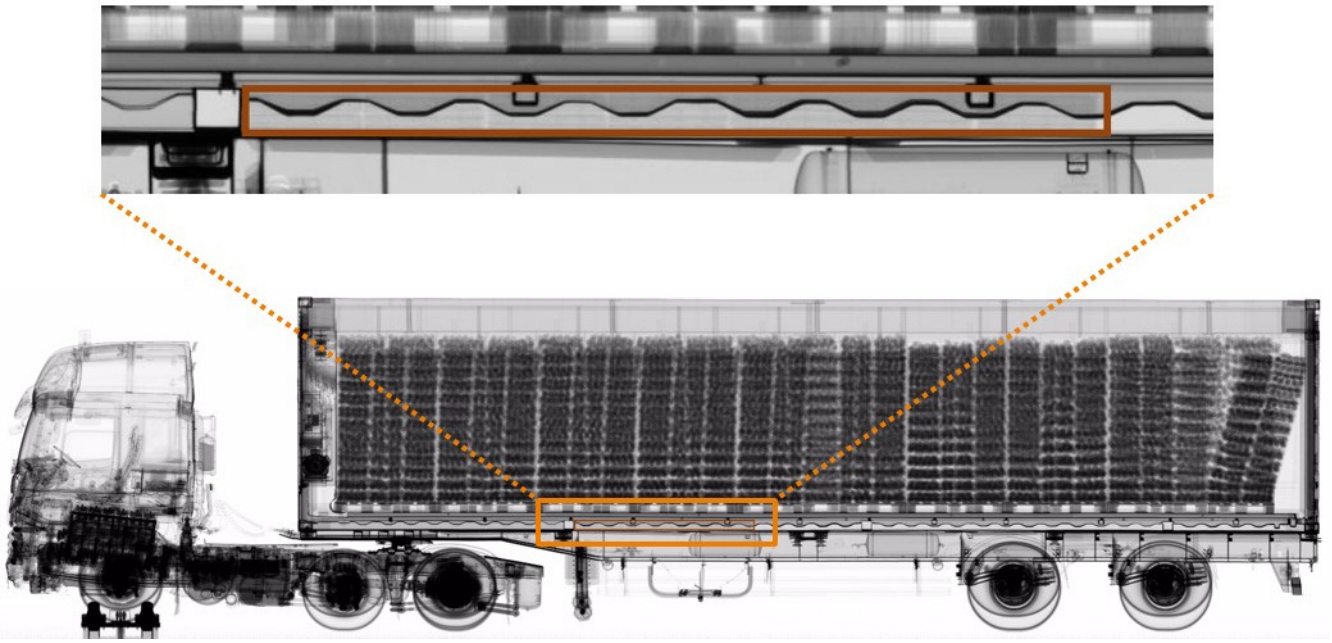


Figure 2: X-ray image of a container with hidden illicit goods

PARTICIPANTS & DESIGN

46 employees of the DTCA and 21 employees of the FCA completed the study. All participants conducted the Ishihara test in order to confirm their color perception ability, and their object recognition skills were assessed with the X-Ray ORT ([8]). Participants that passed the Ishihara test were divided into three groups with comparable object recognition skills (with the exception of the control group of DTCA, which consisted of administrative personnel). As illustrated in Figure 3, the first group did not receive training or assistance by ATRs and served as control group. The second group (training group) was tested and trained without ATRs, while the third group was tested and trained with ATRs. Tests to measure detection performance in cargo screening took place once before training and then again at the end of the study. Each test consisted of two parts, part one with illicit goods only hidden within the freight and part two with half the illicit goods hidden in the container structure (while the other half remained hidden within the freight). In the ATR group (AG), part one used ATRs 1-3 and part two ATR4. Each part consisted of 96 images with illicit goods in one fourth of the images, resulting in eight targets for each of the three categories of illicit goods. For the defined hit and false alarm rate of the ATRs this meant that an alarm indicated an illicit good in 50% of the cases (and missing 25% of the illicit goods). Participants agreed to two training sessions of 15 minutes per week over the course of four months. The training consisted of three difficulty levels where participants progressed to the next level after three to four hours of training.

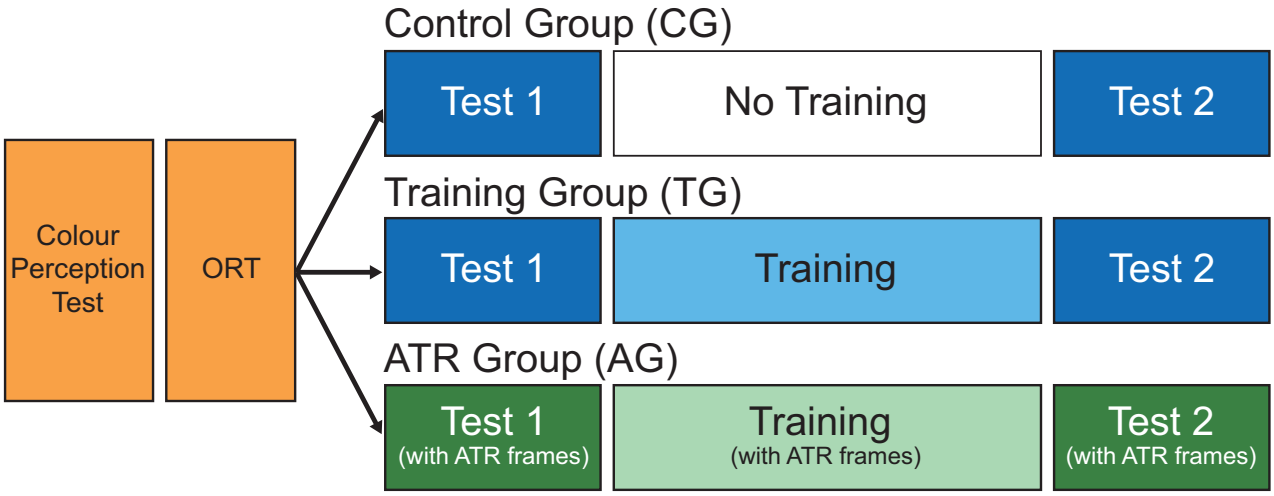


Figure 3: Experimental design of the study

RESULTS & DISCUSSION

Since the tests consisted of images from DTCA only, which differed notably from the images familiar to officers from FCA (resulting in higher false alarm rates), results were analyzed for DTCA and FCA separately and should not be compared between the two customs organizations.

Figure 4 summarizes the results of test part one (illicit goods within the freight only). As can be seen, for the training group detection (hit rate) was higher in Test 2 compared to Test 1, while for the control group, detection remained constant. Statistical analyses confirmed that training significantly improved detection for officers from both customs organizations. Further analyses revealed that this improvement did not differ significantly between the three categories of illicit goods and that the false alarm rate was not affected by training. Figure 4 also shows that detection was higher for the ATR group than the training group, meaning that ATRs for the detection of cigarettes, drugs, and guns could improve detection when looking at these categories combined. Analyses confirmed this improvement to be statistically significant and not to differ significantly between Test 1 and Test 2. This means that the benefit of the ATRs (more specifically: the difference between the training group and ATR group) was about the same before and after training. Further analyses revealed that the false alarm rate was not influenced by ATRs but that the benefit of ATRs differed between the three categories of illicit goods, being larger for the detection of guns than for the detection of cigarettes and narcotics. This might be explained by the size of the guns, which were relatively small and therefore easily missed if not highlighted by the ATR.

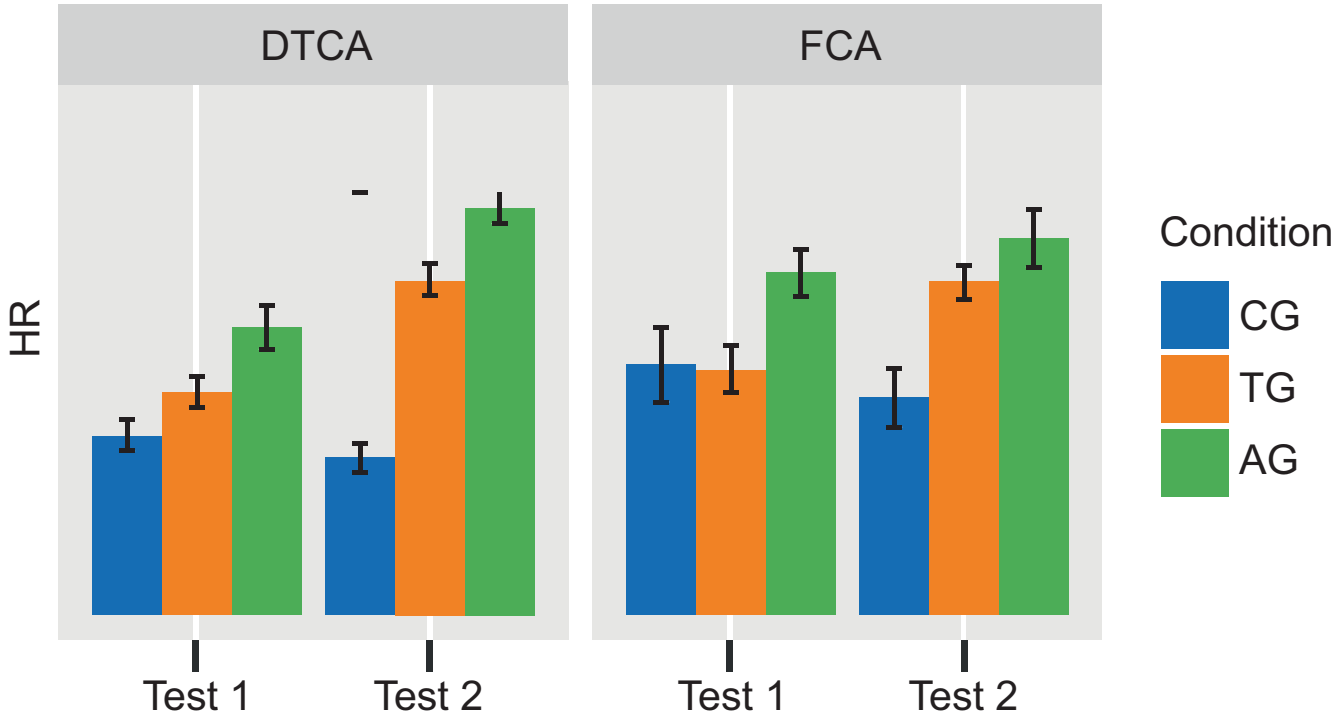


Figure 4: Mean hit rate of test part one, before and after training. Error bars indicate standard errors of the mean

In test part two, where half of the illicit goods were hidden in the container structure (Figure 2) and the other half within the freight, training also had a positive effect on the hit rate without affecting the false alarm rate. But ATR4 (detection of anomalies in the container structure) showed both, benefits and disadvantages: the participants who were assisted by ATR4 detected more illicit goods that were hidden in the container structure, but missed more of the other illicit goods within the freight. A possible explanation for this might be that the ATR directed the officers' attention more towards the container structure and away from the freight.

Overall, the study replicated that training can be an important contribution to improve detection performance in cargo screening. The ATRs led to improved detection performance, but their benefit depended on the category of illicit goods, and ATR4 also showed some undesirable side effects. These results highlight the importance of research in human–machine interaction, before automated support systems are implemented in practice.

Contact us at info@casra.ch to learn more about CASRA's research and applications.

1 While some goods are generally forbidden (e.g. cocaine), many goods require permits and are illicit if not properly declared on the waybill.

2 A hit is defined as a correct alarm by the system, e.g. if a potentially illicit good is hidden in the cargo and the system generates an alarm. If an alarm incorrectly indicates the presence of a potentially illicit good while no such good is hidden in the cargo, this is considered a false alarm.

REFERENCES

[1] Bliss, J., Dunn, M., Fuller, B.S., 1995. Reversal of the cry – wolf effect: An investigation of two methods to increase alarm response rates. Perceptual and Motor Skills, 80 ,1231-1242.

[2] Automation in airport security X-ray screening of cabin baggage: Examining benefits and possible implementations of automated explosives detection. Manuscript submitted for publication.

[3] Halbherr, T., Schwaninger, A., Budgell, G.R., Wales, A., 2013. Airport Security Screener Competency: A Cross-Sectional and Longitudinal Analysis. Int. J. Aviat. Psychol., 23(2), 113-129.

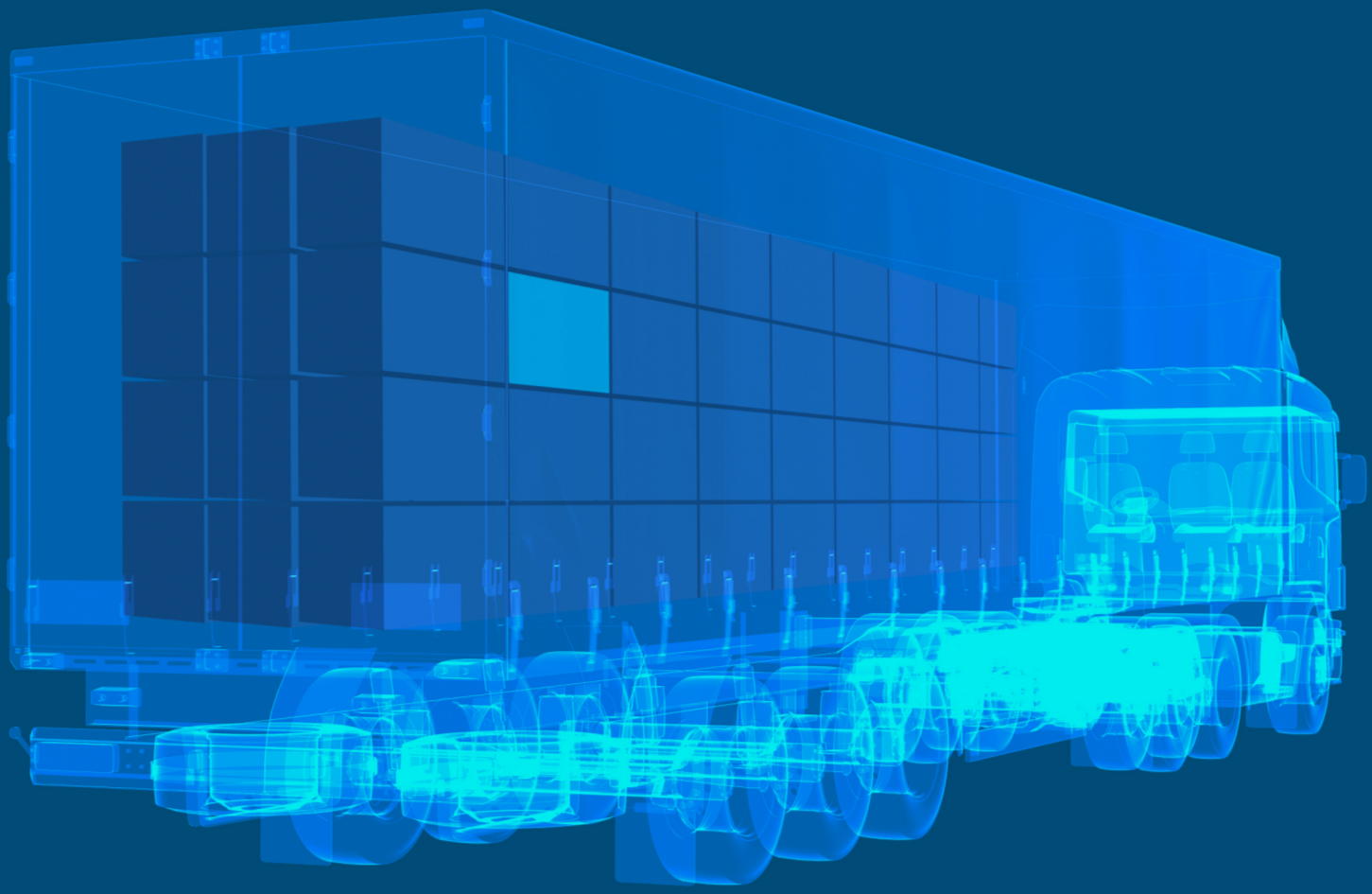
[4] Koller, S., Hardmeier, D., Michel, S., Schwaninger, A., 2007. Investigating training and transfer effects resulting from recurrent CBT of x-ray image interpretation, in D.S. McNamara and J.G. Trafton (Eds.), Proceedings of the 29th Annual Conference of the Cognitive Science Society (pp. 1181-1186). Austin, TX : Cognitive Science Society.

[5] Michel, S., Mendes, M., de Ruiter, J., Koomen, G., & Schwaninger, A. (2014). Increasing X-ray image interpretation competency of cargo security screeners. International Journal of Industrial Ergonomics, 44, 551-560.

[5] Parasuraman R., Sheridan T.B., Wickens C.D., 2000. A model for types and levels of human interaction with automation. IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans. 30(3):286-297

[6] Strauch, B., 2016. The Automation-by-Expertise-by-Training Interaction. Why Automation-Related Accidents Continue to Occur in Sociotechnical Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society. doi: 10.1177/0018720816665459

www.acxis.eu



Empa
Materials Science and Technology

list
cea tech

Fraunhofer
EZRT

Customs Administration of the Netherlands
Tax and Customs Administration

smiths detection
bringing technology to life

CASRA

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössische Zollverwaltung EZV
Administration fédérale des douanes AFD
Amministrazione federale delle dogane AFD