CHAPTER 8

Detection of Irregularities and Abnormal Behaviour in Extreme-Scale Data Streams

Konstantinos Demestichas, Theodoros Alexakis,
Nikolaos Peppes, Konstantina Remoundou,
Ioannis Loumiotis, Wilmuth Muller,
and Konstantinos Avgerinakis

8.1 Introduction

Crime always has been an issue of outmost importance for every society. In 2017, a total of 205 foiled, failed and completed terrorist attacks were reported in the EU [1]. The aforementioned number represents a sharp

K. Demestichas (✉) · T. Alexakis · N. Peppes · K. Remoundou · I. Loumiotis
Institute of Communication and Computer Systems, Athens, Greece
e-mail: cdemest@cn.ntua.gr; talexakis@cn.ntua.gr; npeppes@cn.ntua.gr; kremoundou@cn.ntua.gr; i_loumiotis@cn.ntua.gr

W. Muller
Fraunhofer Institut of Optronics System Technologies and Image Exploitation – IOSB, Karlsruhe, Germany
e-mail: wilmuth.mueller@iosb.fraunhofer.de

K. Avgerinakis
Catalink Ltd, Nikosia, Cyprus
e-mail: koafgeri@catalink.eu

B. Akhgar et al. (eds.), *Technology Development for Security Practitioners*, Security Informatics and Law Enforcement,
https://doi.org/10.1007/978-3-030-69460-9_8

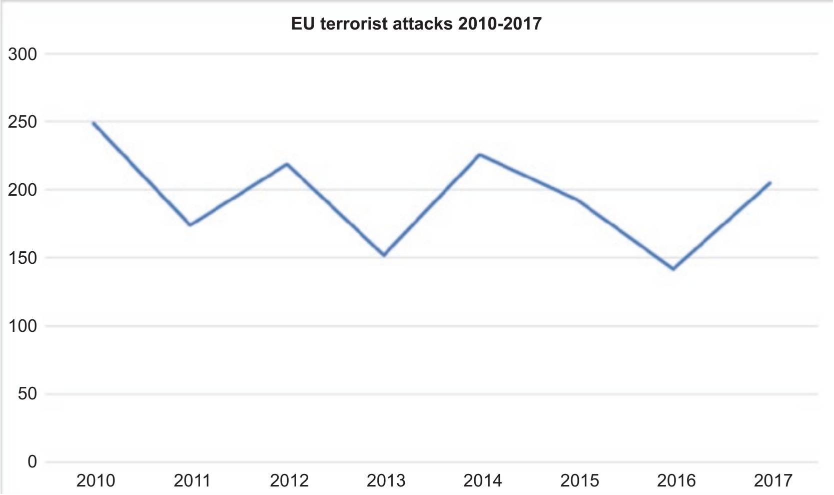

Fig. 8.1 Number of failed, foiled and completed terrorist attacks in the EU 2010–2017

increase around 45% compared to 2016 terrorist attacks and stopped a downward trend which had begun on 2014 (Fig. 8.1).

In the direction of providing increased security, the European Commission developed the European Agenda on Security which sets terrorism, organized crime and cybercrime as interlinked cross-border challenges in which EU countries must have a common strategical approach and coordinated action [2].

Post-study of criminal attacks reveal common patterns among them [3], such as radicalized individuals with history in organized crimes as perpetrators. This underlines the increasing necessity of LEAs to combine, prioritize and analyse heterogeneous massive data streams. Thus, there is the need to integrate sociological, psychological and linguistic models alongside with ICT tools in order to aid predictive policing methods and extract precursors or predictors of abnormal behaviour. The upward trend in cybercrime activities such as attacks on information systems, forms of online fraud and forgery, dissemination of illegal online content and more renders the need even more pertinent. According to cyberterrorism experts, approximately 90 percent of terrorist activity on the Web takes place through the usage of social networking tools [4]. Despite the LEAs' efforts to counter fight terrorism propaganda through social networks, the lack of linguistic capabilities and expertise as well as the differences in assessment of content are being exploited by terrorists in order to infest social media with their outlaw messages [5]. Another relative challenge is that data generation considering crimes is massive and mostly in semi-structured or unstructured data format. Taking into account that it is very difficult to analyse semi-structured or worse unstructured data formats using traditional data mining techniques such as SQL databases and common statistical analysis, Big Data analytics appear as a promising and feasible solution. Additionally, Big Data provide powerful tools by means of social networks analysis, semantic technologies as well as the utilization of advanced linguistic models. Therefore, LEAs must develop and integrate future-proof solutions and tools which will empower them with supreme analytical and predictive capabilities against terrorists, organized crime groups (OCGs) and individuals.

The remainder of the chapter is organized as follows: Sect. 8.2 presents an overview of state-of-the-art research projects in EU; Sect. 8.3 provides a brief review about the available technology; Sect. 8.4 presents the proposed architecture; while Sect. 8.5 concludes the chapter.

8.2 State-of-the-Art Research Projects

The European Union and European Commission realize the necessity of research and development in order to promote state-of-the-art solutions in the field of public safety and security. In this light there are many research programs across Europe most of them totally or partially funded by European Commission. According to the H2020 strategic program, research and state-of-the-art approach is not just about creating new technologies and applying new tools but also requires understanding phenomena such as violent and abnormal behaviour [6]. So, it is clear that alongside with the emerging technologies such as Big Data analytics and machine learning, social and human sciences must be involved. This increases the complexity of all current tools and solutions and makes research mandatory in order to equip LEAs with powerful and future proof tools. Starting from crisis management, the beAWARE project demonstrates an integrated solution to support forecasting, early warnings, transmission and routing of the emergency data as well as the coordination between the first responders and the authorities.

beAWARE aims to provide decision support services to crisis management centres and make first responders and authorities more situational aware. In study [7] of beAWARE project are represented some of the main challenges and solutions such as the collection and integration of heterogeneous data from various sources in a common framework. In addition to crisis management, there are several innovative research projects running with the scope of fighting terrorism and organized crime. The TENSOR project aims to provide LEAs with a terrorism intelligence platform in order to act and plan fast for the prevention and the early detection of organized terrorist activities.

The TENSOR platform integrates a set of tools which allows the detection and gathering of various online data both from the Surface and the Dark Web. In the same direction, the RED-ALERT project brings data mining and predictive analytics tools to the next level, developing novel natural language processing, semantic media analysis, social network analysis, complex event processing and artificial intelligence technologies for online terrorist content. The study [8] is directly connected to the RED-ALERT project and focuses on social media terrorist content and how to remove it. Supplementary to RED-ALERT, the VICTORIA project aims to utilize the data streams from video sources in order to address the need for video analysis for investigation of criminal and terrorist activities. The latest terrorist attacks in London, Brussels, Barcelona, Paris, Berlin and Nice highlight the importance of video recordings. VICTORIA aims to deliver a TRL- 6 Video Analysis Platform (VAP) that accelerates video analysis with reliable results using Big Data tools as described in the paper of Alexander Schindler et al. [9]. One step further is the combination of heterogeneous data streams and the creation of a perpetually self-improving knowledge base according to a sophisticated and representational model which is also the basic idea and concept of MAGNETO project. MAGNETO intends to create a powerful set of tools that will be based on Big Data analytics, semantic reasoning and augmented intelligence well integrated in a common TRL-6 platform. Thus, MAGNETO targets to help LEAs to deal with large volumes and diversity of data in their fight against terrorism and organized crime in general.

Moreover, European Union tries to reduce not only the terrorism rates but also the crime rates generally. In this perspective, the CONNEXIONs project develops and demonstrates next-generation detection, prediction, prevention and investigation based on integration and correlation of multimodal data. CONNEXIONs uses augmented and virtual reality tools in order to construct crime scenes for post or pre-occurrence of a crime. Another major issue the last few years in the EU is the protection of public spaces where large crowds congregate; thus the LETS-CROWD project, which is in collaboration with CONNEXIONs project, aims to overcome the challenges associated with the effective implementation of European Security Models. Study [10] of the LETS- CROWD project focus on human-centred tools and solutions for real-time behaviour forecasting as well as risk assessment methodologies for soft targets and a policy-making toolkit for the long-term and strategic approach of these issues.

Aside from inner security and safety issues, border protection is another one major issue for the EU. Projects like ROBORDER and TRESSPASS aim to provide technologically advanced frameworks which support passenger risk assessment and decision services considering border patrolling and protection. For example, in ROBORDER a fully functional autonomous border surveillance system with unmanned mobile robots and vehicles will be developed which will incorporate multimodal sensors as well as route algorithms for optimal path detection as presented by Athanasios Kapoutsis et al. in [11]. Collected heterogeneous data are analysed and semantically integrated from authorities in order to provide accurate decision support services for border patrolling. TRESSPASS focuses on the utilization of LEA databases in order to contribute to risk-based passenger screening. In the context of border protection, stopping illegal trafficking inside and outside Europe is of high importance. To this end, certain advanced research projects focus on different aspects of illicit markets and illegal trafficking, e.g. the ANITA project. ANITA project and authors of [12] propose a method about textual similarity in video content as a solution which improves the investigation capabilities of LEAs by offering a toolset as well as techniques to efficiently address online illegal trafficking of falsified medicines, NPS, drugs and weapons.

8.3 Available Technologies in Crime Investigations and Future Trends

LEAs, practitioner analysts and investigators have always been equipped with tools and capabilities so they can prevent and fight crime as early as possible. These tools and capabilities are constantly evolving so they can follow and counteract to criminals' advanced methods. In the following paragraphs, we try to make a presentation of the available technologies and their future trends for crime prevention and investigation.

8.3.1 Visual Intelligence

Visual intelligence is one of the most celebrated technologies for crime investigation. It contains several algorithms, leverage deep learning and shallow representation technologies so as to tackle object and person detection, tracking human activity recognition as well as abnormal event detection, face detection and recognition and finally crisis event detection in surveillance and crawled visual data. Object detection and association can be performed by several algorithms such as Fast R-CNN [13], Faster R-CNN [14], YOLO [15], SSD [16], KCF [17], MDNET [18] and VITAL [19]. The same techniques can be used for face detection as well [16, 20], but for face recognition more information is required; therefore, frameworks like DeepFace [21] and FaceNet [22] with more sophisticate deep CNN algorithms are used. Going a step further, vehicle identification uses plate recognition [23] or DCNN [24] features, while activity localization and recognition use activity behaviour patterns of tracked people, objects or vehicles by computing their global spatiotemporal trajectories [25]. Having identification of people objects and vehicles as well as their activity localization and recognition can lead to abnormal behaviour detection which involves trajectory-based analysis on specific tracked objects [26] or statistical-based methods, such as PLSM [27] in order to discover motifs of abnormal activities that may occur in the scene.

A future approach on object detection and tracking can be a hybrid representation of shallow and deep representation features that will encode HOG and SIFT descriptors with a Fisher vector and represent them with a Deep CNN. As far as face detection and recognition are concerned, research should aim to leverage facial point detection and a combination of shallow features with a deep convolutional framework. The object detection representations would also be useful for vehicle detection and representation if combined with a plate DCNN scheme which could lead to a robust vehicle identification technique. Moreover, for action recognition, the goal-based descriptors [28] should be extended with spatiotemporal texture; Fisher vectors and a neural network would then transform these features to apply deep learning capability, while action localization will be deployed with a spatiotemporal saliency detector.

8.3.2 Semantic Integration and Technologies

LEAs and practitioner analysts gather data from a vast variety of sources which in most cases are semi-structured or unstructured. Thus, this information must be transformed to knowledge mainly by cognitive processes conducted by domain experts. In the direction of data and information fusion, the JDL model framework defines how to perform fusion as well as which processes and resources are involved [29]. Today's available tools for semantic reasoning and integration can be divided in three main categories: (i) tools based on mathematical class of random fields, (ii) logical models and frameworks (i.e. if-then rules) and (iii) Markov logic networks. Mathematical class of random fields is a solid and well-understood basis to represent stochastic processes with their random variables and the problem-specific dependencies among them. Logical models represent concepts, instances (objects in the world) and declarative knowledge (rules) in a consistent manner using if-then rules to achieve the fusion. Additionally, the use of OWL (Web Ontology Language) to represent semantic models in an object-oriented manner is a perfect fit for such a logic framework. Markov logic networks basically are the combination of knowledge graphs with stochastic modelling [30] which allows the computation of queries under uncertainties.

The application of information fusion comes with some challenges such as data association, adaptation of the logical model during its usage as well as the need for a high-quality statistical model. Future research should aim to adopt special techniques which incorporate relational knowledge during a probability mapping of the attributes of the related objects in the domain using Big Data analytics tools and technologies. Therefore, future proposals in the field of semantic reasoning and data fusion have to be easily adaptable according to changes in the environment; advanced algorithms, which compute the benefit for incorporating new concepts and rules into the actual model, can realize this.

8.3.3 Data Mining and Detection of Cybercriminal Activities

The World Wide Web is a huge data repository, and LEAs have a huge interest on how to crawl useful data from it. Common approaches focus on the efficiency of the crawling using different methodologies, but they do not take into account the structure of the sources of data. Such common approaches include Focused Crawling [31] or Path Ascending

Crawling [32]. As a counteract to data crawling tools from LEAs, cybercriminals show a rising trend of utilization of various techniques which allow to thwart or at least to delay the detection of their nefarious activities. In order to tackle this and considering that a more flexible and scalable detection solution is needed, data mining-based solutions should be developed that will explore the suitability of pattern discovery, for which initial results seem to be rather promising [33].

Parallel to data mining, security agencies nowadays adopt more and more Big Data technologies in order to apply analytics over gathered information from various online sources such as social media, the Dark Web, etc. [34]. These technologies usually are implemented in the context of i) artificial neural network (ANN) models to predict national security problems in near real time [35], ii) data mining to reveal fraud [36] and iii) classification methods to predict deception in computer-mediated communications [37]. Big Data analytics drawing on online and other activities can be used to determine relevant behavioural indicators. This will help LEAs to detect outliers or anomalies, discover previously unknown associations and rules and continuously monitor data streams as a preventive measure. In this light, the target for future studies, as far as Web intelligence frameworks are concerned, is to leverage existing state-of-the-art data analytics tools, specifically assessing their impact on the information and data stream management.

8.4 Proposed Architecture

This chapter demonstrates the proposed high-level system architecture which provides the required methods for LEAs to accelerate their investigations and remain careful in consideration of terrorist and cybercriminal threats by effectively integrating massive data streams.

The overall strategy of the current proposal comprises an iterative development methodology, where software updates are made available to LEAs and practitioners end-users for testing as well as thorough evaluation and validation purposes. Specifically, they are presented with details relevant to the main functional modules as well as the components that the system generally is consisted of, in a coherent manner. The details supply in a comprehensive way the structure of the system architecture based on the provided requirements together with the combination of the different components into a common proposed platform which will be deployed into LEAs' and practitioners' facilities. Additionally, the initial system-wide functional testing, evaluation and validation will be executed.

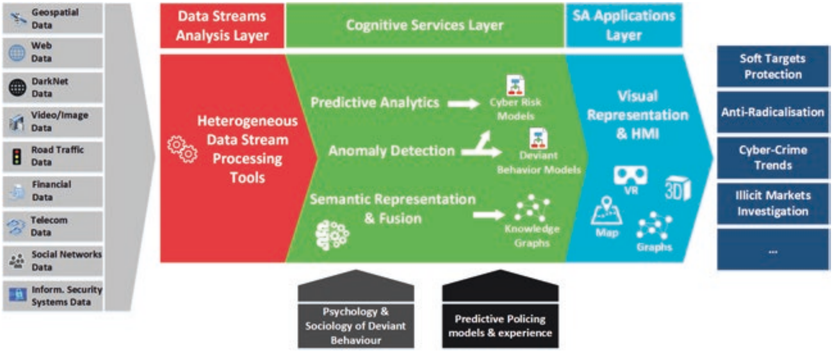

Fig. 8.2 Platform architecture

The main effort is being put on standardization of the platform's open architecture and its constituent components, interfaces as well data exchange formats. In order to succeed, extensive monitoring studying and contribution into activities related to standards of ISO/TC 292 (Security and Resilience) in the area of security is foreseen.

The platform could be considered as a multiple level system with a deeper view of system layers. The system architecture is designed in such a way so as to meet performance and resiliency requirements at scale. The main target of the platform development is to reach TRL-7. The main components of this architecture are shown in Fig. 8.2.

8.4.1 Visual Intelligence Modules

The visual intelligence modules is implemented by using face recognition and face detection algorithms in order to achieve specific identifications of persons contained in images and videos crawled from the Web, social media and/or footage from static or moving cameras. Furthermore, object recognition techniques are also applied in static or wearable surveillance cameras aiming to detect suspicious objects. The choice of the aforementioned techniques could also be extended in order to track people and objects throughout the video. Finally, suspicious or abnormal activities are being tracked by the adoption of crowd analysis and human action recognition with spatiotemporal localization.

The starting point for the prementioned implementations comprises deep convolutional network, DeepFace, SSD coupled with a YOLO architecture, KCF, goal-based descriptors and swarm intelligence for crowd analysis and abnormal activity detection. Therefore, libraries for computer vision and machine learning purposes targeted at dealing with real-time scenarios will be used.

The challenges that arise are data transmission failures and activities or objects that may be occluded. The contingency plan that had to be designed before the experimental implementation contains multiple transmitter and cameras in outdoors pilots, so that possible failures can be tackled.

8.4.2 Data Mining Modules for Crime Prevention and Investigation

Data mining is used to extract valuable information from the existing data. Crawling and mining take place by using crawl points, social media accounts and keywords, from which posts and sites are extracted and stored for analysis purposes. Eventually data from multiple sites are processed and exported into a common format that is available to the other components.

To develop and study crime patterns, we used the existing open-source Web and social media mining components integrated in the TENSOR and HOMER projects in conjunction with APIs provided by social media as well as extensive and scalable open-source Web crawlers software projects. However, Dark Web sites entrance difficulty and specific rate limits on social media access constitute a possible risk. The contingency plan pays attention to key Dark Web Sites and narrow social media access using specific keywords.

8.4.3 Semantic Information Representation and Fusion Modules

A module dedicated to data and information fusion is applied to the heterogeneous data that are collected from different sources, where the use of semantic technologies results to information transformation into valuable knowledge. The baseline of this module includes knowledge graphs, JDL model, OWL language and Markov logic networks along with the use of appropriate semantic information fusion models.

Main problems that arise include the uncertain, various and conflicting data and information, the requirement for solid training data as well the growing distinction between the model and the real world during the model's lifecycle. The use of a manually engineered model for the beginning of the training process, the selection of a confirmed model and the modification of the training frequency are defined as a contingency plan.

8.4.4 *Trend Detection and Probability Prediction Modules for Organized Terrorism and Criminal Activities*

To predict organized terrorism and criminal activities, Big Data analytics techniques are applied over the collected data in order to identify hidden trends inside the datasets. The deployment of analytics results to predictive models is the link between analytics and decision-making process.

Many widely used types of Artificial Intelligence Algorithms such as artificial neural networks, decision trees, pattern recognition and lifelong learning algorithms are developed by using the appropriate libraries as well as suitable predictive policing software. Some problems that occurred are the inaccurate results that were extracted from the prediction model and the false-positive alert that a model generated in some cases. A proposed contingency plan foresees the use of data sources of higher degree of diversity as well the creation of more sophisticated models for explaining deviant behaviours.

8.4.5 *Detection Modules of Cybercriminal Activities*

In addition to identifying anomalies, behavioural indicators and also revealing previously unknown associations and rules that are connected to cybercriminal activities, advanced Big Data analytics techniques, based on artificial neural networks and classification methods, are applied to the collected data.

Three common machine learning algorithms are used, i.e. K-means clustering, support vector machines and deep learning algorithms, and are developed with open-source libraries for numerical computation in order to achieve faster results. Moreover, the suitable ML algorithms for data mining tasks are used.

Risks which appeared in module presented in Sect. 8.4.4 could not only be the inaccurate results of the model, but also model results lose quality over time. The contingency plan gives attention to the selection of more complex model, fit the training frequency as well test and modify the model.

8.4.6 Situation Awareness and HMI Modules

All of the aforementioned functionality has to be demonstrated to the end-users in order to increase the situation awareness of the decision-makers. For this purpose, innovative visualization tools such as virtual and augmented reality technologies are used. The baseline comprises open-source libraries for visual analytics in addition to powerful, secure and flexible end-to-end analytics platforms for data visualization and representation purposes.

Possible problems that may arise and need to be overcome could be the inadequate offered visualization tools for some LEAs, the requirement for additional data views in certain use cases and the difficulty in using and handling the visualization environment. A proposed contingency plan includes the adaptation of the visualization environment, the creation of applicable solutions on top of the existing platform as well more detailed training courseware.

8.5 Conclusions

In this study, we made a brief presentation of state-of-the-art research projects in the domain of crime fight as well as involved technologies and tools. Moreover, we highlighted some of the currently used technologies and tools by LEAs and crime analysts as well as their future prospects. Considering the needs of LEAs for future-proof tools and expertise adoption so to fight and counteract in all types of crimes as well as the increasing resolve of criminals to adopt new technology, we presented a state-of-the-art platform accompanied with a detailed description of the required modules. The modules are integrated and deployed into a common architecture and next will be available to real end-users for detailed validation and evaluation. The architecture comprises all the necessary modules regarding extreme-scale data stream analytics, data mining processes, visual intelligence and machine learning algorithms, semantic and reasoning integration, probability prediction and eventually situation awareness modules and applications, all of them interconnected through open interfaces and a standardized platform. In order to reach TRL-7, an iterative approach will take place with the interaction of real end-users in conjunction with numerous software updates.

Acknowledgements This work has been performed in the context of the PREVISION project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 833115. The chapter reflects only the authors' view, and the Commission is not responsible for any use that may be made of the information it contains.

References

1. Europol. (2018). *TE-SAT 2018: EU terrorism situation and trend report*. Publications Office of the European Union.
2. European Agenda on Security - Migration and Home Affairs - European Commission. (2016). Retrieved from https://ec.europa.eu/home-affairs/what-we-do/policies/european-agenda-security_en.
3. Joint statement of EU Ministers for Justice and Home Affairs and representatives of EU institutions on the terrorist attacks in Brussels on 22 March 2016 - Consilium. Retrieved from http://www.consilium.europa.eu/en/press/pressreleases/2016/03/24/statement-on-terrorist-attacks-in-brussels-on-22-march.
4. Terrorist groups recruiting through social media | CBC News. (2012). Retrieved from https://www.cbc.ca/news/technology/terrorist-groups-recruiting-through-social-media-1.1131053.
5. Europol. Internet Organised Crime Threat Assessment (IOCTA). (2017). Retrieved from https://www.europol.europa.eu/iocta/2017/index.html
6. Horizon 2020 – Work Programme 2018-2020, 14. Secure societies – Protecting freedom and security of Europe and its citizens – European Commission. (2019, July 02). Retrieved from https://ec.europa.eu/programmes/horizon2020/en/h2020-section/secure-societies-–-protect-ing-freedom-and-security-europe-and-its-citizens
7. Hilbring, D., Moßgraber, J., Hertweck, P., Hellmund, T., van der Schaaf, H., Karakostas, A., Kontopoulos, E., Vrochidis, S., Kompatsiaris, I., Gialampoukidis, I., & Andreadis, S. (2018). Harmonizing data collection in an ontology for a risk management platform. In *International conference on informatics for environmental protection (EnviroInfo)* (pp. 126–132). Munich.
8. van der Vegt, I., Gill, P., Macdonald, S., & Kleinber, B. (2019). Shedding light on terrorist and extremist content removal. In *GRNTT*. London: RUSI.
9. Schindler, A., Boyer, M., Lindley, A., Schreiber, D., & Philipp, T. (2019). Large scale audio- visual video analytics platform for forensic investigations of terroristic attacks. In I. Kom-patsiaris, B. Huet, V. Mezaris, C. Gurrin, W. H. Cheng, & S. Vrochidis (Eds.), *MultiMedia modeling. MMM 2019. Lecture notes in computer science* (Vol. 11296). Cham: Springer.
10. Dambra, C., Gralewski, A., & Arias, J. (2019). LETSCROWD: Dynamic risk assessment for mass gatherings. In *Proceedings of the 16th ISCRAM conference*. València, Spain.

11. Kapoutsis, A., Malliou, C., Chatzichristofis, S., & Kosmatopoulos, E. (2017). Continuously informed heuristic A∗-optimal path retrieval inside an unknown environment. In *2017 IEEE Interna- tional symposium on safety, security and rescue robotics (SSRR)* (pp. 216–222). Shanghai.
12. Gkountakos, K., Dimou, A., Papadopoulos, G. T., & Daras, P. (2019). Incorporating textual similarity in video captioning schemes. In *2019 IEEE international conference on engineering, technology and innovation (ICE/ITMC)* (pp. 1–6). Valbonne Sophia-Antipolis, France.
13. Girshick, R. (2015). Fast R-CNN. In *2015 IEEE international conference on computer vision (ICCV)* (pp. 1440–1448). Santiago.
14. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*, 91–99.
15. Hu, P., & Ramanan, D. (2017). Finding tiny faces. In *The IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1522–1530).
16. Derpanis, K. G., Lecce, M., Daniilidis, K., & Wildes, R. (2012). Dynamic scene understanding: The role of orientation features in space and time in scene classification. In *Proceedings CVPR, IEEE computer society conference on computer vision and pattern recognition* (pp. 1306–1313). IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
17. Henriques, J. F., Caseiro, R., Martins, P., & Batista, J. (2015). High-speed tracking with kernelised correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(3), 583–596.
18. Nam, H., & Han, B. (2016). Learning multi-domain convolutional neural networks for visual tracking. In *IEEE conference on computer vision and pattern recognition* (pp. 4293–4302).
19. Song, Y., Ma, C., Wu, X., Gong, L., Bao, L., Zuo, W., Shen, C., Lau, R., & Yang, M. H. (2018). VITAL: VIsual tracking via adversarial learning. In *Proceedings 2018 IEEE/CVF con- ference on computer vision and pattern recognition (CVPR)* (pp. 8990–8999).
20. Jiang, H., & Learned-Miller, E. (2017). Face detection with the faster R-CNN. In *Proceedings 2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017)* (pp. 650–657). Washington, DC.
21. Parkhi, O., Vedaldi, A., & Zisserman, A. (2015). *Deep face recognition*. British Machine Vision Conference (BMVC) 1, pp. 41.1–41.12.
22. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 815–823). Boston.
23. Liu, X., Liu, W., Mei, T., & Ma, H. (2016). A deep learning-based approach to progressive vehicle re-identification for urban surveillance. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer vision – ECCV 2016. ECCV 2016. Lecture notes in computer science* (Vol. 9906). Cham: Springer.

24. Liu, H., Tian, Y., Wang, Y., Pang, L., & Huang, T. (2016). *Deep relative distance learning: Tell the difference between similar vehicles.* 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp. 2167–2175, Las Vegas.
25. Wang, H., & Schmid, C. (2013). *Action recognition with improved trajectories.* 2013 IEEE international conference on computer vision, pp. 3551–3558, Sydney, NSW.
26. Yang, W., Gao, Y., & Cao, L. (2013). TRASMIL: A local anomaly detection framework based on trajectory segmentation and multi-instance learning. *CVIU, 117*(10), 1273–1286.
27. Varadarajan, J., Emonet, R., & Odobez, J. M. (2012). A sequential topic model for mining recurrent activities from long term video logs. *IJCV, 103*(1), 100–126.
28. Tachos, S., Avgerinakis, K., Briassouli, A., & Kompatsiaris, I. (2017). Mining discriminative descriptors for goal-based activity detection. *Computer Vision and Image Understanding*, *160*, 73–86.
29. Liggins, M., II, Hall, D., & Llinas, J. (2008). *Handbook of multisensor data fusion: Theory and practice* (2nd ed.). Boca Raton: CRC Press.
30. Richardson, M., & Domingos, P. (2006). Markov logic networks. *Machine Learning, 62*(1-2), 107–136.
31. Chakrabarti, S., van den Berg, M., & Dom, B. (2000). Focused crawling: A new approach to topic- specific web resource discovery. *Computer Networks, 31*(11–16), 1623–1640.
32. Cothey, V. (2004). Web-crawling reliability. *Journal of the American Society for Information Science and Technology, 55*(14), 1228–1238.
33. Cabaj, K., Mazurczyk, W., Nowakowski, P., & Zorawski, P. (2018). Towards distributed network covert channels detection using data mining-based approach. In *Proceedings of criminal use of information hiding (CUING) workshop co-located with ARES 2018.*
34. Davis, P. K., Walter, L. P., Brown, R. A., Douglas, Y., Parisa, R., & Voorhies, P. (2013). *Using be- havioral indicators to help detect potential violent acts: A review of the science base*. Santa Monica: RAND Corporation.
35. Bueno de Mesquita, B. (2011). Applications of game theory in support of intelligence analysis. In B. Fischoff & C. Chauvin (Eds.), *Intelligence analysis: Behavioral and social scientific foundations* (pp. 57–82). Washington, D.C: National Academies Press.
36. Li, S. H., Yen, D., Lu, W., & Wang, C. (2012). Identifying the signs of fraudulent accounts using data mining techniques. *Computers in Human Behavior, 28*(3), 1002–1013.
37. Zhou, L., Burgoon, J., Twitchell, D., & Qin, T. J., Jr. (2004). A comparison of classification methods for predicting deception in computer-mediated communication. *Journal of Management Information Systems, 20*(4), 139–165.