CHAPTER 15

Border Surveillance Using Computer Vision-Enabled Robotic Swarms for Semantically Enriched Situational Awareness

Georgios Orfanidis, Savvas Apostolidis, Georgios Prountzos, Marina Riga, Athanasios Kapoutsis, Konstantinos Ioannidis, Elias Kosmatopoulos, Stefanos Vrochidis, and Ioannis Kompatsiaris

15.1 Introduction

Political instabilities, war conflicts, economic crises and the maximization of personal profit comprise few of the main causalities that result in increased illegal events at border territories. Cross-border crime is referred to any serious crime with a cross-border dimension committed at or along the external borders [1]. Towards maximizing the overall profit, such activities involve in many cases the utilization of recent technological

G. Orfanidis (✉) · S. Apostolidis · G. Prountzos · M. Riga · A. Kapoutsis
K. Ioannidis · E. Kosmatopoulos · S. Vrochidis · I. Kompatsiaris
Centre for Research and Technology Hellas, Information Technologies Institute, Thessaloniki, Greece
e-mail: g.orfanidis@iti.gr; sapostol@iti.gr; gprountzos@iti.gr; mriga@iti.gr; athakapo@iti.gr; kioannid@iti.gr; kosmatop@iti.gr; stefanos@iti.gr; ikom@iti.gr

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_15

advances such as innovative sensory systems and specialized equipment. Such technological tools facilitate the activities of criminals which eventually might lead even to human casualties as, for example, drug trafficking using unmanned aerial vehicles.

The effective control and identification of transnational crime activities are essential for ensuring peace and stability and for promoting pertinent political and socio-economic activities. At tactical level, European Border Surveillance System (EUROSUR) is a common example for such initiatives. EUROSUR [2] establishes a common framework for the exchange of information and cooperation between EU member states and Frontex to improve situation awareness and reaction capabilities at the external EU borders confronting cross-border crime and protecting lives of migrants. At operational level, considering also the diversity and the increased number of operational aspects, border authorities and relevant practitioners face important challenges in patrolling and protecting areas under their jurisdiction. The heterogeneity of the threats, the wideness of the surveyed areas, the complexity of the operational environments and the adverse weather conditions are some characteristic subjects under consideration from border practitioners. Thus, it is considered imperative in many cases for the operational personnel to be equipped with advanced surveillance systems in order to effectively complete their objectives.

Such systems mostly involve video and thermal cameras; dedicated sensors for motion, pressure, etc.; RFID tags; radars; and satellite images. Despite their sufficient effectiveness, each system displays either environmental restrictions or limited capacities due to spatial heterogeneity. In addition, the majority of these sensory systems are static resulting in restricted monitored areas strictly depending on their technical specifications. As a result, border authorities currently exploit novel technologies posing existing infrastructure as legacy systems. Unmanned vehicles (UxV) provide such cutting-edge technologies that can be utilized as either independent or complement of existing border surveillance equipment. In this book chapter, we introduce and analyse relevant robotic technologies combined with swarm intelligence for a completely autonomous border surveillance system. In addition, pioneer visual detection approaches are presented for increased efficiency, while semantic data representation models upgrade the overall capacities for optimum situation awareness.

The rest of the chapter is organized as follows. Section 15.2 introduces swarm intelligence as an autonomous navigation scheme, while Sect. 15.3 presents enhanced visual detection models. The following section describes semantic enrichment models towards increased situation awareness, while Sect. 15.5 concludes the chapter by highlighting the benefits of such technologies.

15.2 Swarm Intelligence for Autonomous Navigation

The utilization of different UxVs acquires much popularity in missions that demand immediate situation awareness or are considered as hazardous for the integrity of human lives. Due to these technologies, data acquisition from the operational areas of interest is obtained currently safer, faster and more affordable as higher objectives can be accomplished without the need of specialized sensors. However, despite the convenience that a UxV can offer, such systems prerequisite a specialized operator in order to command and manipulate the assets. The complexity of the process is increased in missions where multiple UxVs are commanded to complete one major objective. In such cases, not only the total operator number is increased accordingly, but also the personnel must be in continuous communication to achieve the overall mission.

An autonomous, yet safe and secure, navigation system for operating UxVs has been proven to be essential in numerous application fields. Introducing autonomy for navigation objectives decreases the operator's interference in the overall operations since his involvement from a low-level operator is converted into a manipulator of higher-level objectives for the defined missions, without the requirement of a priori knowledge of utilizing multiple and heterogeneous UxVs. After the identification of high-level objectives, the navigation system will commence to design robot trajectories in order to successfully complete the overall goal of the defined mission. During the execution of the defined mission, the operator acts only as a supervisor nonetheless, for safety reasons; the system is responsive to any interference at any moment. Thus, the process is more effective since the operator can utilize multiple UxVs, without any special expertise and training, while simultaneously, the efficiency of the mission is increased, and the operational time is reduced.

The presented autonomous navigation system, developed specifically for border security operations, supports three different types of missions. More specifically:

- Strictly user-defined paths to be executed separately from UxVs
- Complete coverage of a polygon region of interest (ROI) over a map, utilizing multiple UxVs
- Continuous surveillance of an unknown, dynamically changed ROI utilizing multiple UxVs

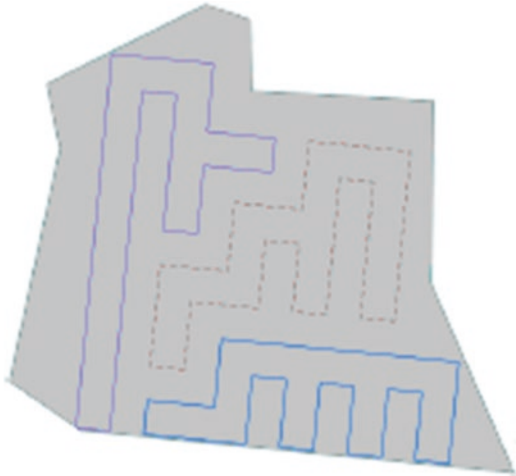
For the first and most simple mission type, the operator/practitioner identifies a set of waypoints for a UxV over a map corresponding to the area of interest. The module provides high-level controls for the UxVs without the need of special training courses or awareness of technical limitations.

Moreover, operating multiple UxVs simultaneously is simplified, while the requirement of using multiple operators is no longer valid. This mission type is considered appropriate for objectives when specific locations must be monitored continuously.

The second type of mission provides the feature of commanding a swarm of UxVs to completely scan a user-defined ROI. Thus, the module is appropriate in covering wide, arbitrary-defined territories benefiting from the number of UxVs in order to significantly limit the overall execution time of the mission and constrain human interference. In addition, it is suitable for different types of UxVs, requiring just minor adjustments on the mission's parameters according to the UxVs' specifications. The overall mission is reduced to a multi-robot Coverage Path Planning (CPP) [3] problem. Receiving as input a polygon for ROI, the number of UxVs and a scanning density (distance between two sequential trajectories), the polygon is represented on an optimized grid for the specific problem, obtaining values that correspond to free space or an obstacle. The entire region is divided into exclusive subregions for every UxV with DARP algorithm [4]. For every subregion, an independent Spanning Tree Coverage (STC) [5] problem is solved. A Minimum Spanning Tree (MST) [6] is constructed, and a circumnavigating path is outlined. These paths incorporate energy aware features, posing them as resource efficient (Fig. 15.1).

Finally, the third mission type provides the capability to the operator to select a region over the map and continuously calculates the optimal monitoring position for every UxV, in order to provide complete situation awareness of the region. The morphology of the region may be completely unknown and dynamically changed, while the number of UxVs may similarly modified even during the mission. The autonomous navigator will

Fig. 15.1 Multi-robot coverage paths in polygon ROI



reallocate the available resources to provide the best possible result and fulfil the overall objective.

A relevant module as reported above was implemented according to a distributed, plug-n-play algorithm for multi-robot applications with a priori non-computable objective functions [7]. This algorithm extracts a sub-cost function individually for each UxV and achieves the overall objective of the swarm by optimizing them combined. Towards this objective, a distributed methodology according to the cognitive-based adaptive optimization (CAO) algorithm [8] is implemented that approximates the evolution of each robot's cost function and adequately optimize its decision variables. The entire training procedure is performed online focusing only on problem-specific characteristics that affect the completion of mission objectives. The fast convergence of the algorithm can ensure fast adaptation of the swarm to the mission, not only during the first stage but also during modifications of the ROI or the swarm itself (Fig. 15.2). As a result, border personnel acting as operators can leverage such systems without requiring specialized training courses, while operational effort is retained at low levels as the feature of autonomy is inherently integrated.

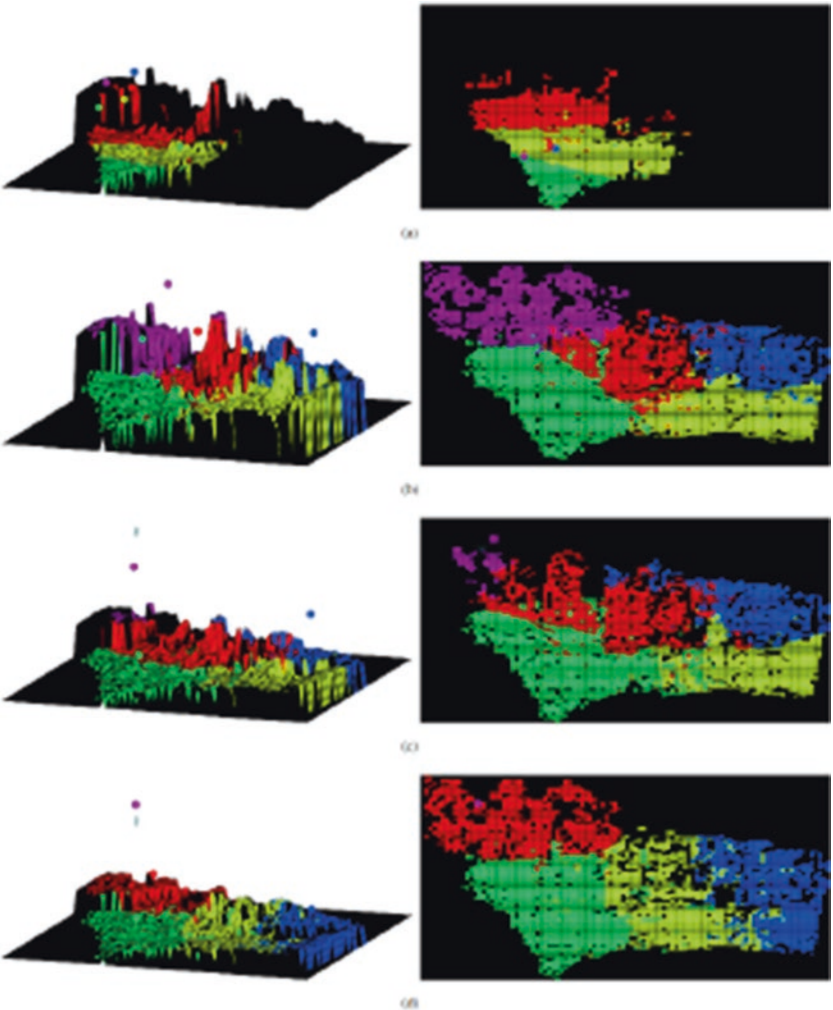


Fig. 15.2 Swarm adaptation to unknown ROI for surveillance during subsequent time steps (**a**–**d**)

15.3 Visual Detection Capabilities

Similarly, due to the heterogeneity of the identified threats, systems utilized by border practitioners should be equipped with enhanced capabilities in identifying specific objects of interest. Considering also that a deployed surveillance system relies on robotic technologies, navigation systems are strictly related to object detection capacities for completeness in the context of autonomous functionalities. In principle, an object detection model corresponds to a schema for simultaneous recognition and localization over the projection plane of objects of interest within a visual representation.

Therefore, the real objective of object detection is to scan the acquired images for identifying any appearance of objects of interest and localizing the detected instances in the processed images. The localization result corresponds to a bounding box surrounding each object of interest, which can be provided in various formats, for example, in upper left and lower right coordinates, centre coordinates, width and height of the bounding box, etc. There are two main categories for visual object detectors: two-step and single-step approaches. The former perform an additional initial step for deciding the "objectiveness" of the area included in a bounding box to determine the best candidates for objects included in the image. The latter category performs both area selection and label assignment (classification) in the same step. The predominant method belonging to the first category is Faster RCNN [9] and typical examples of the second category are Single-Shot Detector (SSD) [10] and You Only Look Once Detector (YOLO) [11] with the latter having several improved versions. The object detector output involves a list of bounding boxes along with their corresponding class labels and their confidence scores. The latter roughly represents the estimation of how confident is the model for the assigned to this bounding box label. Object detection as a capacity is considered overall precise nonetheless, depending on the level of some limitations, inefficient. Thus, a typical approach is to combine this functionality with a tracking module in order to monitor the identified objects. A tracker comprises a module which is provided with an initial bounding box for each detected object and attempts to estimate its motion from a sequence of images or video streams. In most cases, the application of an object tracker is computationally more effective rather than feeding continuously an object detector with sequential frames in systems that require visually identification of specific objects. A typical, yet efficient and fast, tracker relies on the Kernelized Correlation Filters (KCF) [12].

Towards identifying the most efficient object detection model for border surveillance applications, multiple relevant models were deployed and properly evaluated considering both accuracy and execution time. After extensive experiments and evaluations, Faster RCNN [9] resulted in the most sufficient outcomes for the objects of interests as typically, the objects to be identified display small sizes (due to the height and angle of perception) and the model is reported as the most efficient for this objective.

Towards decreasing the overall execution time of the visual identification system, a KCF tracker [12] is applied between two subsequent frames. At every key-frame, an object id is assigned to each distinct object in order to uniquely identify its presence. During the tracking frames, which are typically larger in number than the key-frames, the object ID remains unchanged. At the next key-frames, an Intersection-Over-Union comparison against a fixed threshold of the two bounding boxes is applied. The two bounding boxes, deriving from the object detector and the tracker respectively, are utilized to estimate if the same object is depicted within the bounding boxes' boundaries. The entire scheme is depicted in Fig. 15.3.

For the evaluation process in order to identify the adequacy of the module, the PascalVoc evaluation metric was exploited [13]. The resulted object detection accuracy values are provided in Table 15.1 where 11 classes of objects of interest were used. The presented work emphasized mostly on identifying maritime vehicles leading to identifying four relevant classes: ships, speedboats, inflatable and regular boats. The latter class corresponds to vehicles that could not be categorized in the other classes; nonetheless, the object corresponds to a boat instance. This fact reveals the high importance of maritime border surveillance since the measures

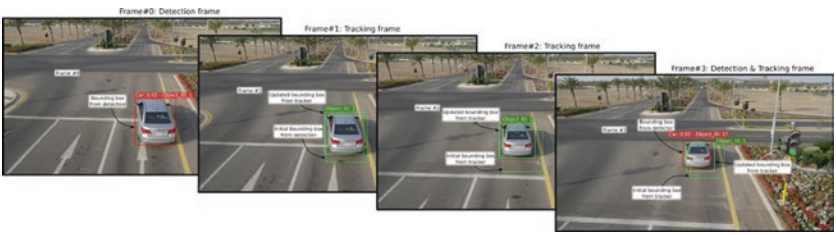


Fig. 15.3 Pipeline for an object tracker in surveillance application

Table 15.1 Accuracy results of Faster-CNN with PascalVoc metric

mAP 0. 6651											
UAV	*Boat*	*Bus*	*Truck*	*Car*	*Helicopter*	*Inflatable*	*Person*	*Motorcycle*	*Ship*	*Speedboat*	
0.68169	0.56543	0.70576	0.64993	0.75568	0.67105	0.41560	0.84015	0.76698	0.73174	0.53311	

that should be considered for each maritime vehicle are diverse; thus, it is imperative to be able to classify such type of vehicles. On the contrary, the performance for some classes suffers since the distinction between these classes is occasionally vague. A typical example of such case would be a light speedboat compared with an inflatable boat with a powerful engine. Figure 15.4 depicts some characteristic examples of visual results acquired with the application of the Faster-RCNN model.

The integration of cognition functionalities comprises a real multilevel asset of the system as object of interests can be identified accurately via processing visual data. Following a hierarchical data flow, the outcomes can be enriched with additional information, while the feature of autonomy can be significantly extended for various operational scenarios.

Therefore, a detailed, yet comprehensive, operational overview can be presented to the operator decreasing the required commanding effort and focusing more on operational goals.

Fig. 15.4 Visual results of Faster-RCNN

15.4 Semantic Enrichment for Increased Situation Awareness

Such surveillance systems display an increased complexity at operational level from the practitioners' perspective as usually, they are not familiarized with such technologies. Noncomprehensive sensor readings and detection outcomes might result in an obsolete system, and eventually, practitioners exploit traditional methods of monitoring the areas of their jurisdiction. In order to facilitate the operational activities of border practitioners and increase their situation awareness, relevant systems integrate technologies at a higher level of implementation to obtain the desired objectives. Such technologies involve the utilization of semantics which refer to the linguistic study of meaning in language coherent to the operator. Therefore, semantic enrichment provides a knowledge framework built upon the acquired data and the detection outcomes so that the operator could be comprehensively be informed.

More specific, ontologies are a means for specifying a vocabulary for conceptualizing and representing a shared domain of discourse [14] in a formal, structured and semantically enriched way. Knowledge in ontologies is modelled via the knowledge graphs by defining common components, like classes (objects, concepts and other entities existing in a domain of interest), properties (attributes, relationships that hold between them), axioms (expressed in a logical form) and rules (if-then statements for logical inferences). With the use of semantic reasoners such as FACT++ [15], Pellet [16] and HermiT [17], logical consequences and new assertions (facts) that are not explicitly expressed in an ontology can be derived.

Ontologies play a key role in facilitating the understanding, sharing and reuse of knowledge between different components within complex systems such as swarm robotics. They have been widely used for situation awareness [18] and decision- making [19] and in IoT infrastructures [20], natural language processing [21] and many more. They demonstrate multiple benefits and capabilities in improved searching, data integration, interoperability, multilinguality and dynamic content generation in an extensive range of areas such as security, healthcare [22], telecommunications, archive portals and law [23].

In the current work, we focus on the semantic representation and enrichment of sensor-based data sourced from different surveillance components (additional sensors, etc.), for extracting potential threats and alerts in the surveillance area, enhancing the representation of the derived detections and improving the situation awareness of the end-users. Eventually, the displayed information to the operator is formatted according to common representation models that are widely utilized in their operational activities at a daily basis.

Therefore, the corresponding service of the increased situation awareness is strictly dependent with the application and the described operational scenarios. More specifically, an ontology was developed for the representation and semantic integration of heterogeneous data generated and exchanged across the cooperative surveillance systems. The proposed semantic model is compliant and extends the EUCISE2020 data model [24], a CISE (Common Information Sharing Environment)-based collaborative initiative for promoting automated information sharing between maritime monitoring authorities. In a nutshell, the CISE data model identifies seven core data entities (Agent, Object, Location, Document, Event, Risk and Period) and eleven auxiliary (Vessel, Cargo, Operational Asset, Person, Organization, Movement, Incident, Anomaly, Action, Unique Identifier and Metadata). An illustration of our ontology-based serialization of the EUCISE2020 model is presented in Fig. 15.5.

The proposed extension of the EUCISE2020 model is related to the following types: (i) further specialization of objects and vehicles and (ii) addition of classes and properties representing the detection of incidents, objects and persons. For demonstration purposes, we consider one rather common scenario in maritime surveillance that involves the detection of an oil spill over sea surface. Whenever an oil spill is detected, an instance of PollutionIncident class (Fig. 15.6) is created, which involves an incident of OilSpill and is associated with respective PollutionType and

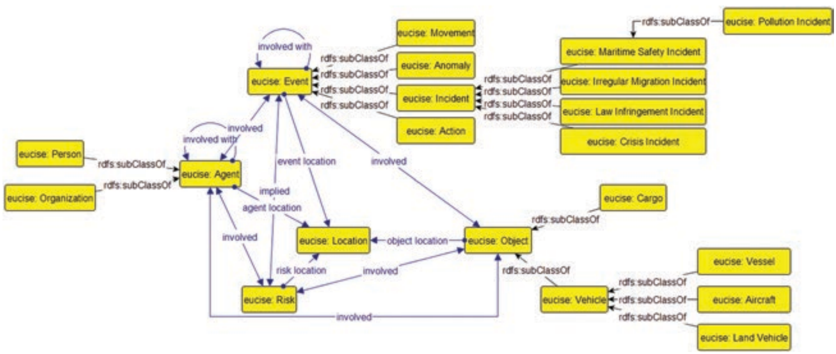


Fig. 15.5 Core classes of our ontology-based serialization of the EUCISE2020 model, along with their main interrelationships

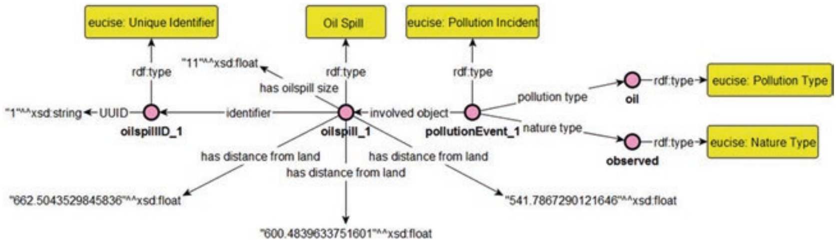


Fig. 15.6 An instance of oil spill associated with a pollution event of specific pollution and nature type

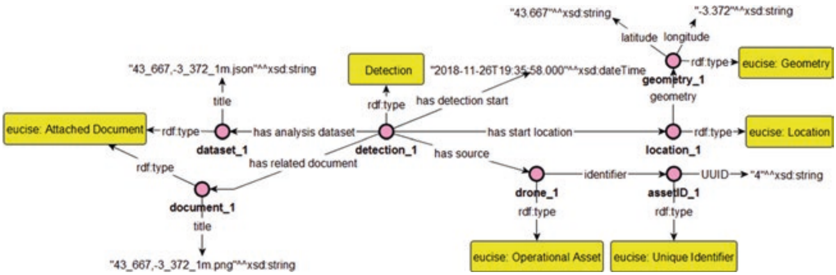


Fig. 15.7 An instance of Detection type associated with an operational asset, a document of reporting and the location of interest

NatureType instances. Also, an instance of Detection is created (Fig. 15.7), which is associated with all relevant information populated in the AttachedDocument, Geometry and the OperationalAsset classes that made the detection via the appropriate data and object properties including hasAnalysisDataset, hasStartLocation and hasSource.

On the basis of the implemented ontology, semantic reasoning techniques (SPARQL rules and constraints) might be additionally adopted to aggregate data from various sources and to achieve both low-level fusion from external resources (such as geospatial services) and high-level fusion by combining information from geographically dispersed and heterogeneous sensors. This approach facilitates the automatic detection and inference of complex events of interest like threats, abnormal activities and illegal border trespassing. In general, SPARQL is a highly expressive RDF query language that allows querying the linked data, by matching one more or patterns against the relationships of the knowledge base while supporting features like aggregation, negation, filtering, constraints and property paths.

Overall, such technologies target eventually to present the system's outcomes within a common representation framework. The displayed alerts and information follow a widely utilized template which was derived from the operational needs of the corresponding experienced personnel. Thus, the system interacts with the operator using one common basis for which the results are comprehensive and intentionally simplified in order for the operators to increase their situation awareness and focus on operational tasks.

15.5 Conclusions

Recent technology advancements are considered to be sufficiently mature for integration in many systems and applications. Even in very complex operational scenarios like border surveillance, cutting-edge technologies can perform adequately well. The relevant practitioners can benefit of such systems towards improving their operational capabilities. As the challenges that they have to confront display significant diversities, the utilized surveillance systems must integrate specialized capacities.

Towards this objective, swarm robotics can broaden the solutions that are provided to the border practitioners. Such systems enhanced with additional features can be used effectively to monitor distant territories. In this chapter, three different pillars of services in different levels of implantation were presented towards describing a fully autonomous and operational surveillance systems. More specific, an optimizer for autonomous navigation of a swarm was presented. The service provides high-level commands to the practitioner to mitigate the complexity of operating such systems while retaining, nonetheless, their effectiveness in monitoring tasks. In addition, visual recognition of object of interests can increase the detection capabilities of the overall system leading to a truly autonomous surveillance framework. Finally, the integration of semantics improve the practitioners' perception for the identifying events increasing the level of the current situation awareness. These three types of technology have been proven particularly efficient in monitoring tasks since they have been extensively deployed in relevant systems as independent features.

Therefore, their integration along with their combination comprises a significant added value for an autonomous surveillance system since each additional feature increases its main operational objective.

Acknowledgements This work was supported by the ROBORDER project funded by the European Commission under Grant Agreement No 740593.

References

1. Passas, N. (2003). Cross-border crime and the interface between legal and illegal actors. *Security Journal, 16*(1), 1 37.
2. Eurosur homepage. https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/border-crossing/eurosur_en. Last accessed: 2 Nov 2019.
3. Cabreira, T., Brisolara, L. R., & Ferreira, P. (2019). Survey on coverage path planning with unmanned aerial vehicles. *Drones, 3*(1), 4.
4. Kapoutsis, A. C., Chatzichristofis, S. A., & Kosmatopoulos, E. B. (2017). Darp: Divide areas algorithm for optimal multi-robot coverage path planning. *Journal of Intelligent & Robotic Systems, 86*(3-4), 663–6 0.
5. Gabriely, Y., & Rimon, E. (2001). Spanning-tree based coverage of continuous areas by a mobile robot. *Annals of Mathematics and Artificial Intelligence, 31*(1-4), 77.
6. Gower, J. C., & Ross, G. J. (1969). Minimum spanning trees and single linkage cluster analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 18*(1), 54–64.
7. Kapoutsis, A. C., Chatzichristofis, S. A., & Kosmatopoulos, E. B. (2019). A distributed, plug- n-play algorithm for multi-robot applications with a priori non-computable objective functions. *The International Journal of Robotics Research, 38*(7), 13–32.
8. Kosmatopoulos, E. B. (2009). An adaptive optimization scheme with satisfactory transient performance. *Automatica, 45*(3), 716–723.
9. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems* (p. 1).
10. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21–37). Springer.
11. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 77–77).
12. Henriques, J. F., Caseiro, R., Martins, P., & Batista, J. (2014). High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(3), 5 3–5 6.
13. Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International Journal of Computer Vision, 88*(2), 303–333.

14. Guber, T. (1993). A translational approach to portable ontologies. *Knowledge Acquisition, 5*(2), 1–22.
15. Tsarkov, D., & Horrocks, I. (2006). Fact++ description logic reasoner: System description. In *International joint conference on automated reasoning* (pp. 2 2–2 7). Springer.
16. Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., & Katz, Y. (2007). Pellet: A practical owl- dl reasoner. *Web Semantics: Science, Services and Agents on the World Wide Web, 5*(2), 51–53.
17. Motik, B., Shearer, R., & Horrocks, I. (2007). Optimized reasoning in description logics using hypertableaux. In *International conference on automated deduction* (p. 67–3). Springer.
18. Frank, A. U. (2001). Tiers of ontology and consistency constraints in geographical information systems. *International Journal of Geographical Information Science, 15*(7), 667–667.
19. Hogenboom, A., Hogenboom, F., Frasincar, F., Schouten, K., & Van Der Meer, O. (2013). Semantics-based information extraction for detecting economic events. *Multimedia Tools and Applications, 64*(1), 27–52.
20. Bimschas, D., Hasemann, H., Hauswirth, M., Karnstedt, M., Kleine, O., Kroller, A., Leggieri, M., Mietz, R., Passant, A., Pfisterer, D., et al. (2011). Semantic-service pro- visioning for the internet of things. Electronic Communications of the EASST 37.
21. Hellmann, S., Lehmann, J., Auer, S., & Brümmer, M. (2013). Integrating nlp using linked data. In *International semantic web conference* (p. 113). Springer.
22. Liyanage, H., Krause, P., & de Lusignan, S. (2015). Using ontologies to improve semantic interoperability in health data. *BMJ Health & Care Informatics, 22*(2), 30315.
23. Benjamins, V. R., Casanovas, P., Breuker, J., & Gangemi, A. (2005). *Law and the semantic web: Legal ontologies, methodologies, legal information retrieval, and applications* (Vol. 336). Berlin/Heidelberg/New York: Springer.
24. Eucise2020 fp7 project. http://www.eucise2020.eu/. Last accessed: 06 Dec 2019.