Border surveillance using computer vision enabled robotic swarms for semantically enriched situational awareness*

George Orfanidis[1], Savvas Apostolidis[1], George Prountzos[1], Marina Riga[1],
Athanasios Kapoutsis[1], Konstantinos Ioannidis[1], Elias Kosmatopoulos[1],
Stefanos Vrochidis[1], and Ioannis Kompatsiaris[1]

Centre for Research and Technology Hellas, Information Technologies Institute,
6th km Harilaou-Thermi, 57001 Thessaloniki, Greece.
`{g.orfanidis,sapostol, gprountzos, mriga, athakapo, kioannid, kosmatop, stefanos, ikom}@iti.gr`

Abstract. Cross-border crime utilizes recent advanced systems to perform their illegal activities. Innovative sensory systems and specialized equipment are examples that were used for illicit human trafficking and for various materials. The increasing challenges that border personnel must resolve require the usage of recent technological advances as well. Thus, the utilization of pioneer technologies seems imperative to precede technologically organized crime. Towards this objective, the introduction of Unmanned Vehicles (UxV) and the advances of relevant sub-systems has created a new solution to fight cross-border crime. Utilizing a combination of UxVs enriched with enhanced detection capabilities comprises an effective solution. The chapter will introduce and present the capability of an autonomous navigation system by exploiting swarm intelligence principles towards simplifying the overall operation. Computer vision advances and semantically enrichment of the acquired information are incorporated to deliver edge-cutting technologies. The described architecture and services can provide a complete solution for optimal border surveillance and increased situation awareness.

Keywords: Border surveillance · Autonomous systems · Visual detection · Semantic representations.

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 to the external borders [20]. Towards maximizing the overall profit, such activities involve in many cases the

---

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

utilization of recent technological advances such as innovative sensory systems and specialized equipment. Such technological tools facilitate the activities of the potential criminals which eventually might lead even to human casualties.

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 to 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 are utilized as complete border surveillance solutions. 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 2 introduces swarm intelligence as an autonomous navigation scheme while Section 3 presents enhanced visual detection models. Following Section describes semantic enrichment models towards increased situation awareness while Section 5 concludes the chapter by highlighting the benefits of such technologies.

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. However, despite the convenience that a UxV can offer, such systems prerequisites 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) [5] 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,

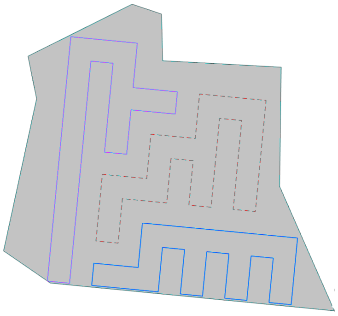

Fig. 1. Multi-robot coverage paths in polygon ROI.

obtaining values that correspond to free space or an obstacle. The entire region is divided into exclusive sub-regions for every UxV with DARP algorithm [14]. For every sub-region, an independent Spanning Tree Coverage (STC) [8] problem is solved. A Minimum Spanning Tree (MST) [9] is constructed and a circumnavigating path is outlined. These paths incorporate energy aware features, posing them as resource efficient (Fig. 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 reallocate the available resources to provide the best possible result and fulfill 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 [15]. 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 [16] 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. 2). As a result, border personnel acting as operators can leverage such systems without requiring specialized training courses.

3 Visual detection capabilities

Similarly, due to the heterogeneity of the identified threats, systems utilized by border practitioners should depict enhanced capabilities in identifying specific

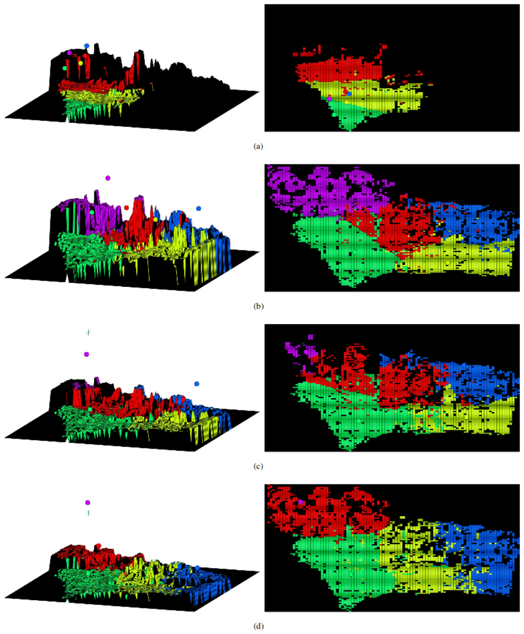

Fig. 2. Swarm adaptation to unknown ROI for surveillance durinjg subsequent time steps (a-d).

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, center coordinates and width and height of the bounding box etc. (Fig. 3). 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 [22] and typical examples of the second category are Single Shot Detector (SSD) [17] and You Only Look Once Detector (YOLO) [21] 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].

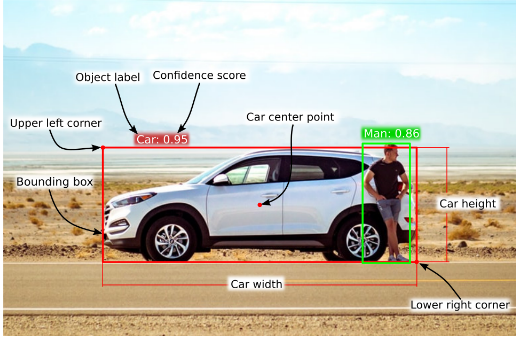

Fig. 3. Example result of a visual detector.

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 [22] resulted 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, resulted from the object detector and from the tracker respectively, are utilized to estimate if the same object is encompassed within the bounding boxes' limitations. The entire scheme is depicted in Fig. 4.

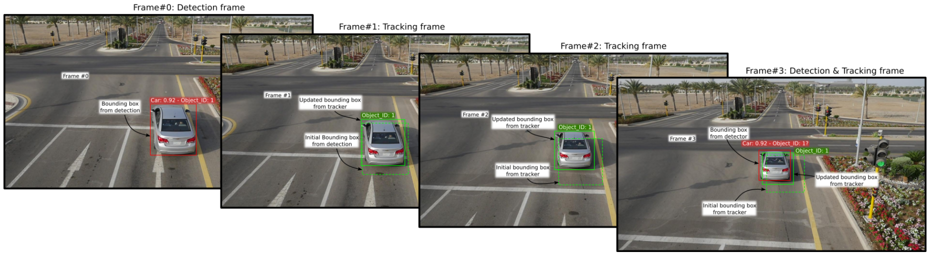

Fig. 4. Pipeline for an object tracker in surveillance application.

Table 1. Accuracy results of Faster-CNN with PascalVoc metric.

mAP 0. 66519										
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

For the evaluation process in order to identify the adequacy of the module, the PascalVoc evaluation metric was exploited [6]. The resulted object detection accuracy values are provided in Table 1 where 11 classes of objects of interest were used. Te presented work emphasized mostly on identifying maritime vehicles leading to identifying 4 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 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 speed-

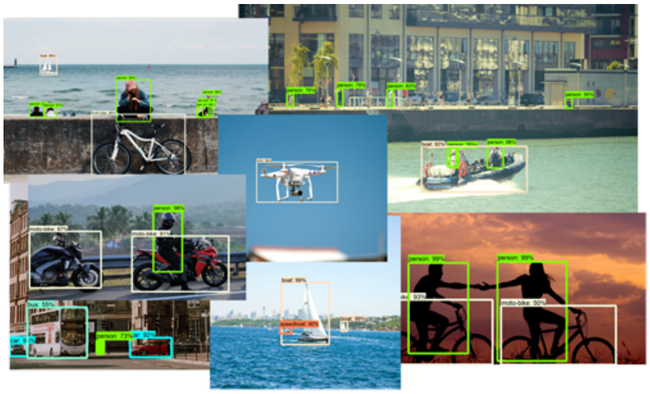

Fig. 5. Visual results of Faster-RCNN.

boat compared with an inflatable boat with a powerful engine. Fig. 5 depicts some characteristic examples of visual results acquired with the application of the Faster-RCNN model.

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 familiarize with such technologies. Non comprehensive sensor readings and detection outcomes might result to 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 involves the utilization of semantics which refer to the linguistic study of meaning in language coherent to the operator.

More specific, ontologies are a means for specifying a vocabulary for conceptualizing and representing a shared domain of discourse [10] 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++ [24], Pellet [23] and HermiT [19], 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 [7], decision-making [13], in IoT infrastructures [4], natural language processing [11] 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 [18], telecommunications, archive portals and law [3]. In the current work, we focus on the semantic representation and enrichment of sensor-based data sourced from different surveillance components (swarm robotics, additional sensors etc.), for extracting potential threats and alerts in the surveillance area, enhancing this way the detections derived from the sensors and improving the situation awareness of the end-users.

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 [1], a CISE(Common Information

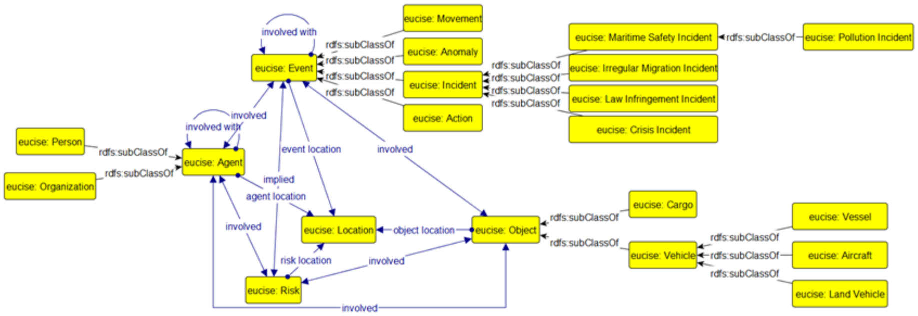

Fig. 6. Core classes of our ontology-based serialization of the EUCISE2020 model, along with their main interrelationships

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 serialisation of the EUCISE2020 model is presented in Fig. 6.

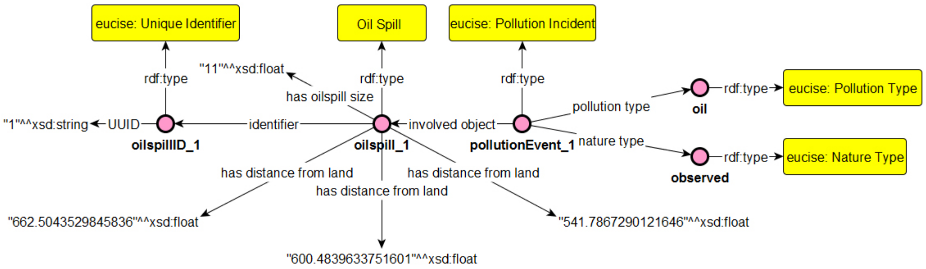

Fig. 7. An instance of oil spill associated with a pollution event of specific pollution and nature type.

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. 7) is created, which involves an incident of `OilSpill` and is associated to respective `PollutionType` and `NatureType` instances. Also, an instance of `Detection` is created (Fig. 8), which is associated with all rele-

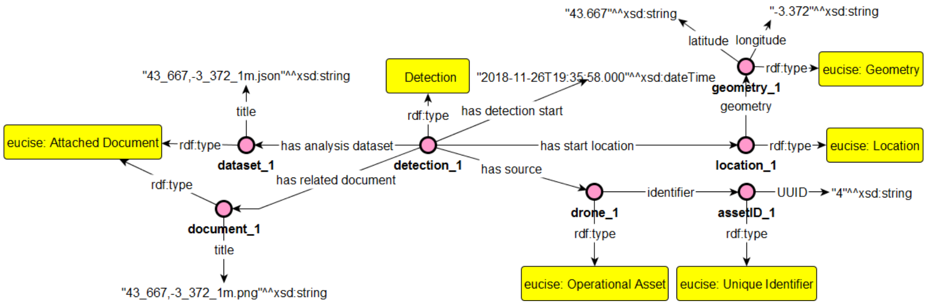

Fig. 8. An instance of Detection type associated with an operational asset, a document of reporting and the location of interest.

vant information populated in the `AttachedDocument`, `Geometry` and the `OperationalAsset` classes that made the detection via the appropriate data and object properties including `hasAnalysisDataset, hasStartLocation, hasSource`.

On the basis of the implemented ontology, semantic reasoning techniques (SPARQL rules and constraints) might be additionalyl 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 details, SPARQL is a highly expressive RDF query language that allows us to query the linked data, by matching one more or patterns against the relationships of the knowledge base, while supporting features like aggregation, negation, filtering, constraints, property paths. In the case of the oil spill detection, we may infer for example if the total size of the oil spill(s) detected is higher than a specific value in cubic meters (m2), or retrieve, for example, the number of entities (persons, vessels) detected close to the oil spill or close to the shore.

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 leading to optimal surveillance solutions.

References

1. Eucise2020 fp7 project. `http://www.eucise2020.eu/`, last accessed: 06 Dec 2019
2. Eurosur homepage. `https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/border-crossing/eurosur_en`, last accessed: 28 Nov 2019
3. Benjamins, V.R., Casanovas, P., Breuker, J., Gangemi, A.: Law and the semantic web: legal ontologies, methodologies, legal information retrieval, and applications, vol. 3369. Springer (2005)
4. Bimschas, D., Hasemann, H., Hauswirth, M., Karnstedt, M., Kleine, O., Kröller, A., Leggieri, M., Mietz, R., Passant, A., Pfisterer, D., et al.: Semantic-service provisioning for the internet of things. Electronic Communications of the EASST **37** (2011)
5. Cabreira, T., Brisolara, L., R Ferreira, P.: Survey on coverage path planning with unmanned aerial vehicles. Drones **3**(1), 4 (2019)
6. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International journal of computer vision **88**(2), 303–338 (2010)
7. Frank, A.U.: Tiers of ontology and consistency constraints in geographical information systems. International Journal of Geographical Information Science **15**(7), 667–678 (2001)
8. Gabriely, Y., Rimon, E.: Spanning-tree based coverage of continuous areas by a mobile robot. Annals of mathematics and artificial intelligence **31**(1-4), 77–98 (2001)
9. Gower, J.C., Ross, G.J.: Minimum spanning trees and single linkage cluster analysis. Journal of the Royal Statistical Society: Series C (Applied Statistics) **18**(1), 54–64 (1969)
10. Guber, T.: A translational approach to portable ontologies. Knowledge Acquisition **5**(2), 199–229 (1993)
11. Hellmann, S., Lehmann, J., Auer, S., Brümmer, M.: Integrating nlp using linked data. In: International semantic web conference. pp. 98–113. Springer (2013)
12. Henriques, J.F., Caseiro, R., Martins, P., Batista, J.: High-speed tracking with kernelized correlation filters. IEEE transactions on pattern analysis and machine intelligence **37**(3), 583–596 (2014)

13. Hogenboom, A., Hogenboom, F., Frasincar, F., Schouten, K., Van Der Meer, O.: Semantics-based information extraction for detecting economic events. Multimedia Tools and Applications **64**(1), 27–52 (2013)
14. Kapoutsis, A.C., Chatzichristofis, S.A., Kosmatopoulos, E.B.: Darp: Divide areas algorithm for optimal multi-robot coverage path planning. Journal of Intelligent & Robotic Systems **86**(3-4), 663–680 (2017)
15. Kapoutsis, A.C., Chatzichristofis, S.A., Kosmatopoulos, E.B.: 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), 813–832 (2019)
16. Kosmatopoulos, E.B.: An adaptive optimization scheme with satisfactory transient performance. Automatica **45**(3), 716–723 (2009)
17. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C.: Ssd: Single shot multibox detector. In: European conference on computer vision. pp. 21–37. Springer (2016)
18. Liyanage, H., Krause, P., de Lusignan, S.: Using ontologies to improve semantic interoperability in health data. BMJ Health & Care Informatics **22**(2), 309–315 (2015)
19. Motik, B., Shearer, R., Horrocks, I.: Optimized reasoning in description logics using hypertableaux. In: International Conference on Automated Deduction. pp. 67–83. Springer (2007)
20. Passas, N.: Cross-border crime and the interface between legal and illegal actors. Security Journal **16**(1), 19–37 (2003)
21. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 779–788 (2016)
22. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. pp. 91–99 (2015)
23. Sirin, E., Parsia, B., Grau, B.C., Kalyanpur, A., Katz, Y.: Pellet: A practical owl-dl reasoner. Web Semantics: science, services and agents on the World Wide Web **5**(2), 51–53 (2007)
24. Tsarkov, D., Horrocks, I.: Fact++ description logic reasoner: System description. In: International joint conference on automated reasoning. pp. 292–297. Springer (2006)