Advancing Frontiers in SLAM: A Survey of Symbolic Representation and Human-Machine Teaming in Environmental Mapping

Brandon C. Colelough, iD

School of Engineering and Information Technology, University of New South Wales, Australia

Abstract - **This survey paper presents a comprehensive overview of the latest advancements in the field of Simultaneous Localization and Mapping (SLAM) with a focus on the integration of symbolic representation of environment features. The paper synthesizes research trends in multi-agent systems (MAS) and human-machine teaming, highlighting their applications in both symbolic and sub-symbolic SLAM tasks. The survey emphasizes the evolution and significance of ontological designs and symbolic reasoning in creating sophisticated 2D and 3D maps of various environments. Central to this review is the exploration of different architectural approaches in SLAM, with a particular interest in the functionalities and applications of edge and control agent architectures in MAS settings. This study acknowledges the growing demand for enhanced human-machine collaboration in mapping tasks and examines how these collaborative efforts improve the accuracy and efficiency of environmental mapping**

I. INTRODUCTION

As described in [1], there is an opportunity to use "artificial creativity to generate sensible narratives that summarise complex military situations for an operator that could enhance situational awareness (SA) where that operator is not already immersed in the low-level tactical situation". Simultaneous Localisation and Mapping (SLAM) offers a solution to this capability gap and has great applicability over a range of military domains (see the appendix for use cases) and for which many systems have already demonstrated promising results [2], [3], [4], [5] (see anything within the semantic-SLAM circle of figure 1 for more). The ethical use of AI-enabled systems within the military is predicated on the ability of the operator to effectively command such systems, as "command is a fundamentally human function that cannot be conducted by machines; it provides accountability for ethical decision making" [6]. The Defence Science and Technology Group list trust as a key facet required in any AI-enabled system to ensure an acceptable level of command is maintained over these systems [7]. The transparency in an AI-enabled system is the "operator's awareness of an autonomous agent's actions, decisions, behaviours and intentions" [8]. This system transparency is "an essential element to facilitate effective communication and collaboration" and as such is the "overarching concept" required for a human-machined teamed environment trust architecture [9]. For further elaboration on the ethics of using AI-enabled systems within the Australian Defence Force, see the video listed here [10]. Recently, structured knowledge graphs known as ontologies have been applied to the field of simultaneous localisation and mapping (SLAM) (see figure 1 - Semantic SLAM, of note is [5], [11], [12], [13], [14], [15]) allowing for semantic reasoning to generate more effective mapping techniques. Generally, two main orientations for contextual reasoning are employed, which are centralised and collaborative. Centralised observes agents transfer information to a back-end system for processing, whereas collaborative contains agents with the capability to locally process information and reason collaboratively [16]. Recent technological advances in artificial intelligence (AI) for visual reasoning have made it possible for edge agents to more effectively reason about their local environment [17] leading to a more effective integration of symbolic and sub-symbolic reasoning. A domain for which symbolic representation has not yet been fully realised, however, is the use of symbolic reasoning through ontological design to create contextual maps of an environment.

II. LITERATURE REVIEW

A systematic literature review (SLR) is the process of "identifying, evaluating and interpreting all available research relevant to a particular research question or topic area or phenomenon of interest" [127]. To prove the capability gap for the proposed system existed, a systematic literature review surrounding the core concepts of the thesis main and sub questions was conducted in accordance with the PRISMA review standard [128] and with consideration to the software engineering SLR guidelines listed in [127].

A. Search Process

A manual search process was conducted to identify published peer-reviewed articles. This process utilised the following databases:

1) IEEE Explore: http://ieeexplore.iee.org
2) Scopus: https://www.scopus.com/
3) Semantic Scholar: https://www.semanticscholar.org/
4) Google Scholar: http://scholar.google.com
5) arxiv online library: https://arxiv.org/
6) Science Direct: www.sciencedirect.com
7) Multidisciplinary Digital Publishing Institute (MDPI): https://www.mdpi.com/
8) The Association for Computing Machinery (ACM) digital: http://portal.acm.org/portal.cfm
9) SpringerLink Library: www.springerlink.com

In addition to the papers selected from these databases, the bibliographies of papers selected through the SLR process were scanned for additional papers (snowballing). The search terms used to conduct the SLR were:

1) multi-agent AND SLAM
2) semantic AND SLAM
3) symbolic AND SLAM
4) ontology AND SLAM
5) human AND swarm AND teaming
6) human AND MAS AND teaming

B. Inclusion / Exclusion Criteria and Quality Assessment

Articles were included if they were:

- Published in a conference or journal or were book chapters
- Articles published from 2010 to 2022
- Full text articles

Articles were excluded if they were:
- Not written in English
- Similar articles from different databases
- Articles not relating to the main concept areas being SLAM or MAS or symbolic reasoning

C. Paper Selection

Five focal areas directly applicable to the main research question and the mechanisms to achieve it were formulated. These include:

1) Semantic SLAM
2) ontologies for SLAM
3) Multi-Agent Systems
4) Simulation
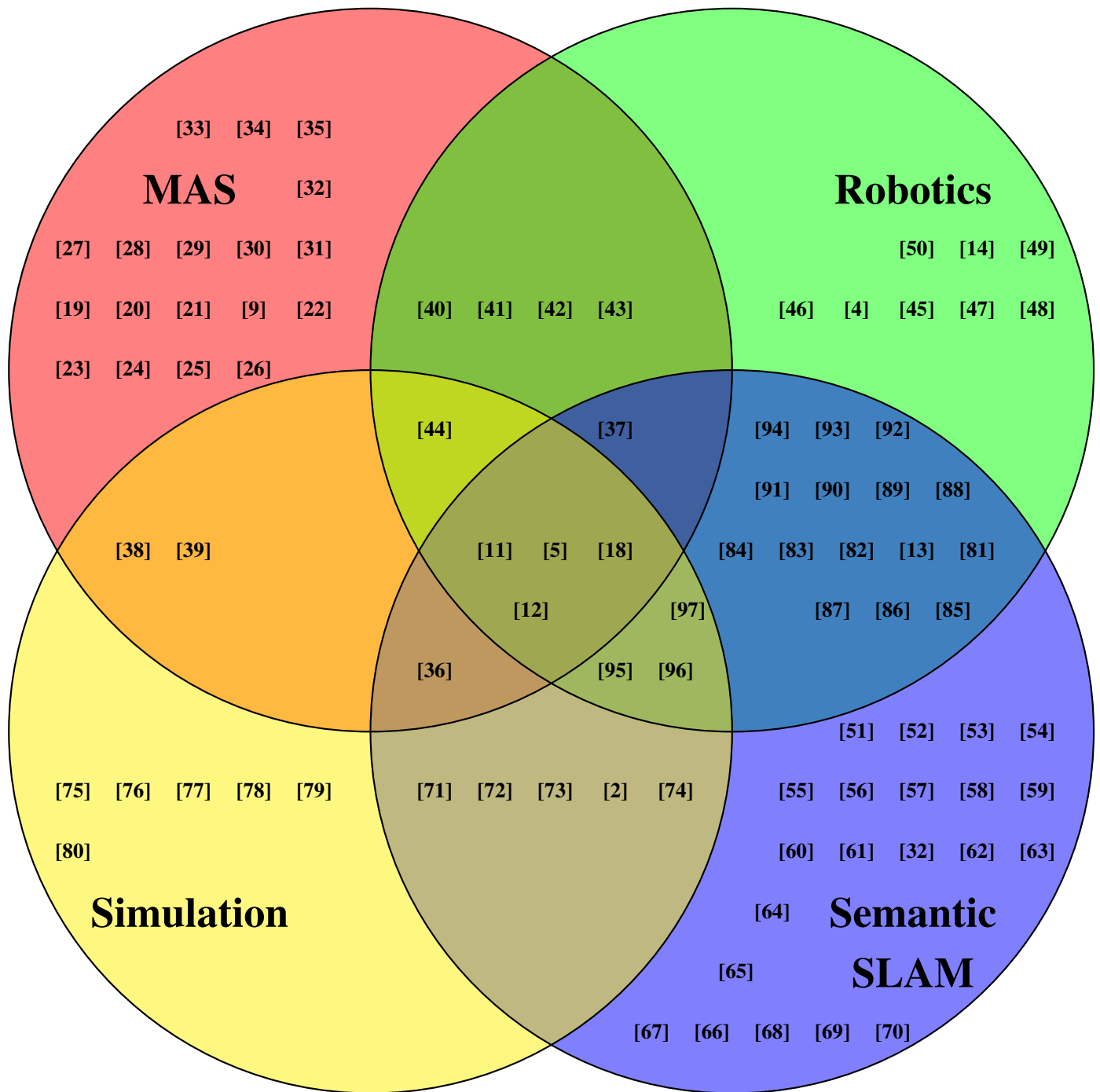
Fig. 1: A literature review of existing architectures surrounding the concept of the proposed SYMBO-SLAM architecture was conducted. This literature review focused using symbolic reasoning through ontological design to create contextual maps of an environment. The Venn diagram in the centre links the majority of concepts required for the implementation of the proposed system in simulation and on hardware. The box below lists the papers directly applicable to the use of symbolic reasoning in the SLAM domain

5) Robotics

Nine additional areas were also considered during the literature review process that formulated key areas of research applicable to this project, which included:

1) ontologies and Symbolic Reasoning
2) Map Matching
3) Localisation
4) 2D Mapping Techniques
5) 3D Mapping Techniques
6) Agent Search and Edge Reasoning
7) Path / Mission Planning
8) Tracking and Spatio-Temporal Reasoning
9) Semantic Segmentation

The papers obtained from the manual search were then evaluated against these areas for relevance. The title, abstract and conclusion were considered to determine whether the paper related to the main concept areas for the project as listed above. 117 papers were then selected that had direct relevance to the five focal areas for the research project. These papers were then reviewed in depth and categorised to be listed in the Venn diagram shown in figure 1. This diagram aims to illustrate the specific area of relevance for each article reviewed in the systematic literature review, and links the majority of concepts required for the implementation of the proposed system in simulation and on hardware. These 117 papers were then also sub-categorised further to show relevance to each of the nine additional areas of relevance listed above. During this process, the snowball method was utilised to highlight an additional

Literature Results from Database queries						
Database	multi-agent SLAM	semantic SLAM	symbolic SLAM	ontology SLAM	human swarm team	human MAS team
IEEE	28	23	4	6	9	1
Scopus	33	30	35	23	37	8
Semantic Scholar	30	30	37	35	35	31
Google Scholar	38	30	2	2	50	0
ArXiv	11	30	1	0	17	3
Science Direct	4	26	0	0	4	0
MDPI	7	30	0	2	4	0
ACM	2	9	0	0	1	0
Springer	2	18	0	0	4	0

TABLE I: The 6 search terms were queried through the 9 databases as described above. The number of pieces of literature returned from each query is shown in the table above. Note also that results that returned above 30 literature pieces were further filtered by year published for reduction

78 pieces of literature from the references in the original set that had direct relevance to subject areas surrounding the use of symbolic-reasoning through ontological design to create contextual maps of an environment on a MAS with SLAM. The 195 articles in total were then placed in the categorised boxes surrounding the Venn diagram. These sub-categories aimed to support the underlying theory of the proposed system. See the appendix - figure 23 for the Venn diagram figure of the full literature review conducted.

D. Agent search and path planning

It is natural to assume that the optimal path will be the main area of interest for search and detection problems. This is often not the case, and instead the distribution effort of searchers (especially on a MAS) is the object of interest for effective search strategies. As Washburn describes in his book [129], there are two good reasons for this:

1) Most searchers would not benefit from knowing the optimal path unless it happened to be easy to follow
2) a "path" is not a convenient mathematical object in most cases

As such, Washburn poses that for a fixed region, a random search pattern will yield a detection rate in less time than that of an exhaustive search. This decrease in detection time does however come at the cost of area coverage. With a strategy attempting to cover the searchable fraction $A' = VWt$ of area A (where V is the speed and, W is the sweep width of a region and t is the time taken), an exhaustive search will cover A'/A of that area (or else all of that area). This is not the case however for a random search strategy. Washburn uses the analogy of a searcher randomly dropping confetti within the area A to demonstrate this, stating that a random search is "effectively confetti casting, and it achieves a lower detection probability than the exhaustive search because of the wasteful overlap of one piece of confetti on another". Resource management in a constrained networked environment is vital for distributed multi-agent based systems such as the one proposed in this project. Al-Asfoor et al. [25] propose search algorithms based on a multi-agent model for dynamic and heterogeneous networks. This search algorithm makes use of conventional random walk algorithm with a semantic inspired resource search paradigm to direct search messages based on the semantic closeness of a target object. Duncan et al. [24] further introduce coverage and targeting Levy strategies that are biologically inspired and have been adapted to robotic swarm systems, which take their basis again in random search strategies.

E. Feature extraction

The term feature extraction in machine learning typically refers to the process of dimensionality reduction by which an initial set of raw data is reduced to a more manageable group for processing. This process aims to find the most compacted and informative set of features (distinct patterns) to enhance the efficiency of a classifier [130]. We define feature extraction as "the process of extracting key features (landmarks) from the environment through visual processing," which is specific to the setting proposed. Current object detection architectures such as EfficientDet [131], ASSFF [132], CenterMask [133] and ATSS [132] offer an average precision between 43% and 50% depending on a multitude of variables such as speed and hardware implementation. Presently, the well recognised YOLO algorithm dominates the field of object detection with its fourth

implementation [134] (note that the fourth version is the latest from the original creator and with a published paper) and comes packaged with pre-trained weights and an enormity of training data-sets (ImageNet, VOC, Kaggle etc).

F. Place Recognition

Place recognition is the ability to identify a feature regardless of change in pose (viewpoint) or other factors of appearance within a global map. This technique is commonly used in literature to enable feature localisation on a global co-ordinate system. In their 2009 survey, Williams et al. [135] compared the state-of-the-art place recognition techniques used for loop closure technology. They concluded that, due to the sparsity of information, map matching techniques under-performed when compared to image matching techniques. ORB-SLAM [136] [137] is a well known visual-SLAM architecture which utilises image matching techniques for loop closure detection. The embedded place recognition module was based on the bag-of-words place recognition technique [138] and implements a covisibility graph of features known as ORBs -Orientated FAST (key point detector architecture [139]) and Rotated BRIEF (de-scriptor architecture [140]). These features are binary features invariant to rotation and scale [141] used to simultaneously conduct tracking, local mapping and loop closing. This system dominated the V-SLAM field for a period post its release in 2015 prior to being superseded by more updated machine-learning place recognition techniques. A more recent study from 2021 [142] compared the deep learning approaches to place recognition, which utilised methods on the spectrum of supervised to unsupervised learning categories. End-to-end frameworks used to address a domain translation problem for place recognition were also explored in this survey, for which NetVlad [143] was found to have had the largest impact. In the domain of supervised learning, three approaches were explored, being holistic, landmark and region based. From this, landmark based supervised training was found to effectively solve for appearance change, perceptual aliasing and viewpoint changing. These techniques (as demonstrated in [144], [145], [146], [147], [148]) employ feature extractors (Generally CNN's or similar) to develop semantic-based feature labels used to identify potential landmarks in a visual feed (see localisation on figure 23 for more). Several systems (such as [2], [33], [149] to name a few) utilise these extracted land-marks to generate a pose-graph of key features within an environment. As such, the area of pose graph optimisation is a relevant field of study for land-mark based SLAM (see [150] for a good overview on this).

G. Multi-Agent System

An agent within a MAS has characteristics of self-sufficiency, social capacity and reactivity [151]. A multi-agent system consists of any number of agents greater than two. Processing on a Distributed Artificial Intelligence (DAI) network can be done either through Mobile Cloud Computing (MCC) or Multi-Access Edge Computing (MEC). The MCC approach observes a system where the majority of the processing requirements is completed at the back end of the system, whereas MEC employs an architecture where the majority of the processing is done at the edge [19] [20]. Similarly, two main orientations for the contextual reasoning of an environment through edge agents are generally given as centralised contextual reasoning where agents transfer information to a back-end system for processing

and collaborative where agents locally process information and reason in a collaborative manner [16] [21].

H. Map Matching

All SLAM algorithms require some form of map merging technique to be applied to enable the continuous updating of the known environment currently being mapped. As described in [135], three main approaches were traditionally used, being map-to-map, image-to-image and image-to-map. These approaches generally took the form of frame-to-frame, single rigid body transformations to relate reference frames in different maps. Early implementations of SLAM such as [3] would generally use spatial matching between sub-maps to achieve the map matching process. This was done by representing the 6D voxels of an RGB-D sensor or point-cloud system as a Gaussian Distribution that can subsequently be integrated together as a map-merging technique. [31] details a networked solution for feature merging through map alignment and data association for applications with SLAM solutions on a MAS. Similar map matching techniques are still being advanced today as shown in [14] where semantic labelling is integrated with geometric map matching to generate a more effective SLAM result for loop closure. SLAM techniques that utilised key frame solutions allowed for map matching to occur as a data insertion into a graph rather than the geometric matching techniques used previously as seen here [137] [2], [149] with pose-graphs. Of interest was a technique proposed in [27] that allowed for cross-platform map merging between platforms that did not share the same sub-mapping techniques, which proposed the use of semantic representation for common understanding. For a more in depth explanation on traditional map matching methods, see [142] sub-paragraph 3 – belief generation. Recently, techniques for map matching utilising landmark features (as described above) have seen success. [146] demonstrates a ConvNet landmark-based visual place recognition system that utilises sequence search and hashing-based landmark indexing, which greatly increased the efficiency of the map-matching process. [144] is another place recognition architecture that employs semantic understanding of landmark features for an effective map matching. This system employed CNN-based key-point matching, which utilised semantic filtering and dense descriptor weighting to allow for a place search procedure leading to a candidate match selection function.

I. Ontology and Symbolic Reasoning

Symbolic AI is the term used for a number of related AI methods that attempt to reason about problems using high-level human understandable representations (symbols) [7]. For contrast, sub-symbolic AI systems (such as artificial neural networks, deep learning etc.) attempt to learn operations from (often large) data sources to map the variables in a problem space.

$$f : X \rightarrow Y$$

A key takeaway from this is that (even at high levels of complexity), sub-symbolic methods for machine learning determine some output Y through a learned function f (e.g., weights in a CNN) without having a clearly explainable or transparent process for obtaining said outputs. As such, some form of symbolic reasoning is required to generate a level of trust in the production of useful ground maps capable of summarising complex situations. This would enable their authenticity and as such enhance the situational awareness of an operator that is not already immersed in the low-level tactical situation [7]. As described in [152], an ontology is a set of entities, which can be classes, properties, or individuals. ontologies may be used to standardise the knowledge base of a specific domain and allow readability by both human and machines, as has been demonstrated in [22]. Ontologies are designed to represent complex knowledge sets about things and the relations between them. As described in chapter 1 subsections 2-3 of [153], the common components of an ontology are its individuals (or instances, I), classes (or concepts, C), attributes (a) and the relationships (R) linking them. Restrictions and rules can be placed upon these components to assert some knowledge about

them. Axioms (A) are developed within the ontology to generate assertions to describe the overall theory of the ontology for its application domain. These assertions are provided in a logical form, such as for example:

$$A \implies B$$
$$A \text{ exists}$$
$$\therefore B$$

And hence, an ontology can be described (as it commonly is in literature) as a five-tuple [154]:

$$O = < C, R, a, I, A >$$

Transparency refers to an operator's awareness of an autonomous agent's actions, decisions, behaviours and intention [1]. This system transparency is the overarching concept required to enable a trust architecture within the human-machine teamed environment. This transparency can be achieved through multiple methods such as interpretability, explainability and predictability [9]. Utilising these ontologies provides a method to share semantic knowledge, promoting bi-directional transparency between humans and artificial agents [154] through the three key facets listed above. Therefore, the sub-set of symbolic AI that is ontological design offers enables explainability, interpretability and predictability within a system which is required to achieve transparency in an artificially intelligent system and hence enable the teamed operator to trust a swarm of robots to complete a cooperative function [155]. ontologies have been applied extensively in many fields to incorporate expert level domain knowledge into symbolic solutions. A corpus for ontology and semantics was proposed in [56] that leverages the textual data documentations for semantic labelling and modelling from 101 separate data sets, which presents a great opportunity for researchers to create more thorough ontological-domain representative models. Utilising these hierarchical ontologies for guided learning in sub-symbolic systems is a concept that is explored by Campbell [44] (Chapter 6). This application enables a sub-symbolic system to reason on abstract concepts and reduce the dimensionality of a problem space (through partitioning) by applying prior knowledge to a learning system. Such a system is described by Hepworth in [44] (Chapter 7). Here, a hybrid approach to activity recognition is detailed that fuses both data- and ontology-driven approaches to the activity recognition problem space in the machine learning domain. A similar approach to symbolic reasoning through the use of a knowledge graph to allow for a more diverse object detection system is shown in [156] and a similar use of semantics to broaden the knowledge depth of scenes for VQA tasks is shown in [157]. The RoboEarth framework presented in [158], [159] proposed a system that inherently integrates a knowledge base with visual SLAM, which in turn allowed for more accurate representation of the environment and recognition of the objects found within. The semantic modelling framework Triplet Ontological Semantic Model (TOSM) proposed in [5] utilised short and long term memory with a static ontology to create of an on-demand ontological knowledge graph for the representation of an environment in the ontological domain (see figure 1 - ontologies for SLAM for more).

The HST-3 architecture as described in [9] was developed to enable transparency for the application of human-machine teaming. The authors here specify that interoperability, explainability and predictability are the tenets required to enable transparency in an AI system which allowed for the operator to further trust the system as a whole. The decentralised nature of reasoning conducted on a MAS is described here as a main challenge in explain-ability (and hence transparency) of a system. The accumulation of varying agent experiences leads to a heterogeneous knowledge set, which hinders convergence to a central solution. To alleviate this issue, control agents are utilised to ensure the responsibility of explanation rests on only a few agents. Similarly, the proposed SYMBO-SLAM architecture aims to use the control agent architecture to alleviate this issue. The HST-GO architecture built upon the HST-3 framework described in [22] aims to achieve transparency in the human-machine interaction space through bi-directional information flow. The tenets of lexicon, syntax and semantics

are used here to describe the shared language between the individuals of the system and further enable communication in the human-swarm teaming environment. The most advanced ontology designed specifically for SLAM tasks is presented in [15]. This ontology suite standardises the SLAM problem set, combining data from a range of ontology sets to achieve superiority at the domain knowledge, lexical and structural levels. The ontology provides knowledge on the four areas of robot information, environment mapping, time information and workspace information. The environment mapping knowledge set includes functionality to describe areas using both geographical and landmark based positioning. Particularly useful in this ontology is the inclusion of uncertainty in robot and landmark positions, which assists greatly in modelling the dynamics of a SLAM problem. Although this ontology currently dominates the SLAM field (see figure 1 - ontology for others) it does not possess the knowledge set required for multi-agent systems. The onto4MAT ontology however is the first attempt to design an ontology for multi-agent teaming systematically [154]. This ontology enables an operator to provide an intent as tasks to a multi-agent system, and for the agents to then provide feedback to the operator.

J. Sub-symbolic and Symbolic SLAM in Dynamic Environments

Spatial-temporal reasoning (as defined by the APA dictionary of Psychology) is "the ability to conceptualise the three dimensional relationships of objects in space and and to mentally manipulate them as a succession of transformations over a period of time." This ability to conceptualise a 3D environment is vital for pose estimation (and as such environmental mapping). For SLAM methods that use image-to-image based tracking methods, some sort of spatial-temporal reasoning is required to maintain a map of the environment. Methods similar to that proposed in [137] use ORBs as described above to construct feature points to achieve this. This architecture sees the key frames tracked in a spanning tree, which allows for the direct mapping of map points on a co-visibility graph. However, this SLAM technique often fails to accurately map environments in dynamic scenarios that contain moving objects (e.g., people). Motion tracking is the ability of a system to be able to follow the motion of an object across multiple frames of a video. Deep Sort (in conjunction with YOLO) is one of many algorithms currently available that provides this capability [160]. Motion tracking of dynamic objects in a scene has recently been combined with filtering techniques and Pose-Slam technology to alleviate the dynamic scenario problem alluded to above [55], [57], [59], [51], [52], [53], [54], [161], [162], [163]. The common theme shared amongst all of these papers is that YOLO is used to detect objects and a low-level semantic understanding is used to determine whether they are dynamic. This is then used to determine whether to filter out parts of a frame, which is then fed into OrbSlam2 to allow for environmental mapping. The issue that plagues sub-symbolic SLAM algorithms in dynamic environments is less of a concern however, in architectures that utilise symbolic methods. Utilising landmark detection or similar with either preprocessing or an embedded module ensures that the system is much more resilient to dynamic feature interference. This is because the semantic label associated with each feature allows for the system to understand whether each feature is dynamic or static inherently.

K. Simulated Environment

For use in this project, we assert that a simulated environment as a computer-based 3D platform capable of supporting a number of different sensor types launched on simulated robotic platforms to produce close-to-reality motion patterns. Many environments were considered for the deployment of the proposed architectures such as the DCIST env as shown in [73], the Gazebo env [80], the Webots env [79] and many others (see simulation on Venn diagram in figure 1 for more). The EyeSim env [164] was selected for its hardware interoperability and support network through both documented resources and ongoing development support networks. This environment supports the deployment of many simulated robotic platforms in a joint space with visual, depth-based and IMU based sensor feeds. Agent communication is supported through radio and Wi-Fi based communications platforms, and the supported development languages are python and C / C++. The Unity-3D physics engine is the basis for the EyeSim environment, which incorporates error into the robotic movements. The intended hardware platforms (EyeBots) are described in [164] and the Unity-based system simulation environment aims to provide a realistic platform to match the actual EyeBot movements.

L. Robotics

The MAS SYMBO-SLAM architecture is to be deployed on both simulated and physical robotic platforms for experimental and testing purposes. The implementation on hardware will focus on unmanned ground vehicle (UGV) platforms capable of ground search tasks. [3], [40] and [14] utilise wheeled systems to achieve SLAM on a MAS. Other implementations include a system mounted to a car as seen in [165] or [86] and the use of aerial drones to conduct ground mapping observed in [150] or target localisation shown in [30]. Combinations and extensions of these implementation strategies, such as synchronous UAV and UGV deployment [47] are also possible (see figure 1 - robotics for more implementation methods).

III. CONCLUSION

This survey paper has extensively explored the integration of symbolic representation in Simultaneous Localization and Mapping (SLAM) within the context of multi-agent systems (MAS) and human-machine teaming. The systematic literature review, conducted following the PRISMA standard and software engineering guidelines, has identified a significant corpus of research that bridges the gap between symbolic reasoning, ontological design, and their application in SLAM tasks. The review highlights the burgeoning interest in deploying symbolic reasoning through ontological design to create more effective, context-aware maps in both centralized and collaborative agent systems. Advancements in artificial intelligence, particularly in edge computing and visual reasoning, have catalyzed the integration of symbolic and sub-symbolic reasoning, offering promising avenues for enhanced environmental mapping and situational awareness. This survey underscores the criticality of transparency in AI-enabled systems, especially in military applications, where the ethical use of such technology hinges on the operator's ability to command and trust the system. The fusion of ontological knowledge with existing SLAM techniques presents a promising direction for future research, especially in creating dynamic, context-rich maps that enhance operator awareness in complex environments. The deployment of these advanced systems in simulated and real-world scenarios further demonstrates their potential applicability across various domains. This paper emphasizes the importance of continued exploration in this field, particularly in enhancing the trust architecture within human-machine teamed environments through improved system transparency. The potential of symbolic reasoning in SLAM, combined with the advancements in MAS and human-machine collaboration, paves the way for more robust, efficient, and transparent mapping systems, crucial for a wide range of applications, from military operations to autonomous navigation.

REFERENCES

[1] D. S. Kate Devitt, Michael Gan, "A method for ethical ai in defence, defence science and technology group," 2020.

[2] A. Rosinol, M. Abate, Y. Chang, and L. Carlone, "Kimera: From slam to spatial perception with 3d dynamic scene graphs,," *Int. J. Robot. Res*, 2021, doi: arXiv:2101.06894.

[3] R. Reid and T. Bräunl, "Large-scale multi-robot mapping in magic 2010," in *2011 IEEE 5th International Conference on Robotics, Automation and Mechatronics (RAM)*, 2011, pp. 239–244.

[4] Y. Yue, C. Zhao, R. Li, C. Yang, J. Zhang, M. Wen, Y. Wang, and D. Wang, "A hierarchical framework for collaborative probabilistic semantic mapping," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 9659–9665.

[5] S.-H. Joo, S. Manzoor, Y. G. Rocha, S.-H. Bae, K.-H. Lee, T.-Y. Kuc, and M. Kim, "Autonomous navigation framework for intelligent robots based on a semantic environment modeling," *Applied Sciences*, vol. 10, no. 9, 2020, doi: 10.3390/app10093219. [Online]. Available: https://www.mdpi.com/2076-3417/10/9/3219

[6] J. W. Council, "Adf concept for command and control of the future force," 2019.

[7] G. Moy, S. Shekh, M. Oxenham *et al.*, "Recent advances in artificial intelligence and their impact on defence," Defence Science and Technology Group, 2020.

[8] G. Moy, S. Shekh, and M. O. et al., "Recent advances in artificial intelligence and their impact on defence," Defence Science and Technology Group, 2020.

[9] A. J. Hepworth, D. P. Baxter, A. Hussein, K. J. Yaxley, E. Debie, and H. A. Abbass, "Human-swarm-teaming transparency and trust architecture," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 7, p. 1281–1295, Jul 2021.

[10] B. Colelough, "How can artificial intelligence ethically be employed in warfare," 2022. [Online]. Available: https://youtu.be/-wH7mUJnFIY

[11] Y. Tian, Y. Chang, F. H. Arias, C. Nieto-Granda, J. P. How, and L. Carlone, "Kimera-multi: Robust, distributed, dense metric-semantic slam for multi-robot systems," *IEEE Transactions on Robotics*, pp. 1–17, 2022, doi: 10.1109/TRO.2021.3137751.

[12] I. Kostavelis and A. Gasteratos, "Semantic mapping for mobile robotics tasks: A survey," *Robotics Auton. Syst.*, vol. 66, pp. 86–103, 2015.

[13] K. J. Doherty, D. P. Baxter, E. Schneeweiss, and J. J. Leonard, "Probabilistic data association via mixture models for robust semantic slam," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 1098–1104.

[14] Y. Yue, M. Wen, C. Zhao, Y. Wang, and D. Wang, "Cosem: Collaborative semantic map matching framework for autonomous robots," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 4, pp. 3843–3853, 2022.

[15] M. A. Cornejo-Lupa, Y. Cardinale, R. Ticona-Herrera, D. Barrios-Aranibar, M. Andrade, and J. Diaz-Amado, "Ontoslam: An ontology for representing location and simultaneous mapping information for autonomous robots," *Robotics*, vol. 10, no. 4, 2021, doi: 10.3390/robotics10040125. [Online]. Available: https://www.mdpi.com/2218-6581/10/4/125

[16] C. Anagnostopoulos and K. Kolomvatsos, "Predictive intelligence to the edge through approximate collaborative context reasoning."

[17] C. Zhang, J. Yu, Y. Song, and W. Cai, "Exploiting edge-oriented reasoning for 3d point-based scene graph analysis," *CVPR*, 2021.

[18] M. Waibel, M. Beetz, R. D'Andrea, R. Janssen, M. Tenorth, J. Civera, J. Elfring, D. Gálvez-López, K. Häussermann, J. M. M. Montiel, A. C. Perzylo, B. Schießle, O. Zweigle, and R. van de Molengraft, "Roboearth - a real world wide web for robots," in *ICRA 2011*, 2011.

[19] P. Madduru, "Artificial intelligence as a service in distributed multi access edge computing on 5g extracting data using iot and including ar/vr for real-time reporting," *IT in Industry*, vol. 9, no. 1, 2021.

[20] L. Y. Jie Jiang, J. Jiang, Y. Liu, and B. Cui, "Angel: a new large-scale machine learning system," *National science review*, 2017.

[21] H. L. J. van der Maas, A. O. Savi, A. Hofman, and K.-J. K. M. M. *The, "The network approach to general intelligence," *General and Speci c Mental Abilities*, 2019.

[22] D. P. Baxter, A. J. Hepworth, K. F. Joiner, and H. Abbass, "On the premise of a swarm guidance ontology for human-swarm teaming," in *65th Annual Meeting of the Human Factors and Ergonomics Society*, Baltimore, Maryland, 2021.

[23] A. J. Hepworth, K. J. Yaxley, D. P. Baxter, K. F. Joiner, and H. Abbass, "Tracking footprints in a swarm: Information-theoretic and spatial centre of influence measures," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2020, pp. 2217–2224.

[24] S. Duncany, G. Estrada-Rodriguez, J. Stocek, M. Dragone, P. A. Vargas, and H. Gimperlein, "Efficient quantitative assessment of robot swarms: coverage and targeting levy strategies," 2022.

[25] M. F. Muntasir Al-Asfoor, "A study of multi agent based resource search algorithms," 2012.

[26] S. P. Yadav, D. P. Mahato, and N. T. D. Linh, "Distributed artificial intelligence," 2020.

[27] I. Andersone, ""heterogeneous map merging: State of the art," *Robot*, vol. 8, no. 3, p. 74–103, 2019.

[28] A. Wiktor and S. Rock, "Collaborative multi-robot localization in natural terrain[*]," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 4529–4535.

[29] S. Saeedi, L. Paull, M. Trentini, M. Seto, and H. Li, "Group mapping: A topological approach to map merging for multiple robots," *IEEE Robotics Automation Magazine*, vol. 21, no. 2, pp. 60–72, 2014.

[30] F. Gu, Y. He, and J. Han, "Active persistent localization of a three-dimensional moving target under set-membership uncertainty description through cooperation of multiple mobile robots," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 8, pp. 4958–4971, 2015.

[31] S. Yu, C. Fu, A. K. Gostar, and M. Hu, "A review on map-merging methods for typical map types in multiple-ground-robot slam solutions," *Sensors (Basel, Switzerland)*, vol. 20, 2020.

[32] V. Tchuiev and V. Indelman, "Distributed consistent multi-robot semantic localization and mapping," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4649–4656, 2020, doi: 10.1109/LRA.2020.3003275.

[33] P. Lajoie, B. Ramtoula, Y. Chang, L. Carlone, and G. Beltrame, "Doorslam: Distributed, online, and outlier resilient slam for robotic teams," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, p. 1656–1663, 2020, doi: https://doi.org/10.1109/LRA.2020.2967681.

[34] M. J. Schuster, K. Schmid, C. Brand, and M. Beetz, "Distributed stereo vision-based 6d localization and mapping for multi-robot teams," in *Journal of Field Robotics*, vol. 36, no. 2, 2019, pp. 305–332.

[35] P. Schmuck and M. Chli, "Ccm-slam: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams," in *J. Field Robot*, vol. 36, no. 4, 2018, p. 763–781.

[36] S.-H. Joo, S. Manzoor, Y. G. Rocha, H.-U. Lee, and T.-Y. Kuc, "A realtime autonomous robot navigation framework for human like high-level interaction and task planning in global dynamic environment," *ArXiv*, vol. abs/1905.12942, 2019.

[37] W. Benny, J. Kun, Y. Mengmeng, W. Tuopu, T. Xuewei, and Y. Diange, "Crowdsourced road semantics mapping based on pixel-wise confidence level," *Automotive Innovation*, vol. 5, 2022, doi: 10.1007/s42154-021-00173-x. [Online]. Available: https://doi.org/10.1007/s42154-021-00173-x

[38] Y. Wang, Y. Yue, M. Shan, L. He, and D. Wang, "Formation reconstruction and trajectory replanning for multi-uav patrol," *IEEE/ASME Transactions on Mechatronics*, vol. 26, no. 2, pp. 719–729, 2021.

[39] D. Zou and P. Tan, "Coslam: Collaborative visual slam in dynamic environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 354–366, 2013, doi: 10.1109/TPAMI.2012.104.

[40] M. J. Schuster, C. Brand, H. Hirschmüller, M. Suppa, and M. Beetz, "Multi-robot 6d graph slam connecting decoupled local reference filters," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 5093–5100.

[41] R. Reid, A. Cann, C. Meiklejohn, L. Poli, A. Boeing, and T. Braunl, "Cooperative multi-robot navigation, exploration, mapping and object detection with ros," in *2013 IEEE Intelligent Vehicles Symposium (IV)*, 2013, pp. 1083–1088.

[42] S. Saeedi, M. Trentini, M. Seto, and H. Li, "Multiple-robot simultaneous localization and mapping:a review," *Journal of Field Robotics*, 2015.

[43] J. Fan, P. Zheng, and S. Li, "Vision-based holistic scene understanding towards proactive human–robot collaboration," *Robotics and Computer-Integrated Manufacturing*, vol. 75, p. 102304, 2022, doi: https://doi.org/10.1016/j.rcim.2021.102304. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584521001848

[44] H. Abbass and R. Hunjet, *SHEPHERDING UXVS FOR HUMAN-SWARM TEAMING*, 1st ed. SPRINGER NATURE, 2022.

[45] J. Fan, P. Zheng, and S. Li, "Vision-based holistic scene understanding towards proactive human–robot collaboration," *Robotics and Computer-Integrated Manufacturing*, vol. 75, p. 102304, 2022, doi: https://doi.org/10.1016/j.rcim.2021.102304. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584521001848

[46] Q. Zhang, M. Wang, and Y. Yue, "Robust semantic map matching algorithm based on probabilistic registration model," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 5289–5295.

[47] Y. Yue, C. Zhao, Z. Wu, C. Yang, Y. Wang, and D. Wang, "Collaborative semantic understanding and mapping framework for autonomous systems," *IEEE/ASME Transactions on Mechatronics*, vol. 26, no. 2, pp. 978–989, 2021.

[48] Y. Yue, C. Zhao, M. Wen, Z. Wu, and D. Wang, "Collaborative semantic perception and relative localization based on map matching," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 6188–6193.

[49] Y. Yue, P. G. C. N. Senarathne, C. Yang, J. Zhang, M. Wen, and D. Wang, "Hierarchical probabilistic fusion framework for matching and merging of 3-d occupancy maps," *IEEE Sensors Journal*, vol. 18, no. 21, pp. 8933–8949, 2018.

[50] Y. Yue, C. Yang, Y. Wang, P. G. C. N. Senarathne, J. Zhang, M. Wen, and D. Wang, "A multilevel fusion system for multirobot 3-d mapping using heterogeneous sensors," *IEEE Systems Journal*, vol. 14, no. 1, pp. 1341–1352, 2020.

[51] C. Yu, Z. Liu, X.-J. Liu, F. Xie, Y. Yang, Q. Wei, and Q. Fei, "Ds-slam: A semantic visual slam towards dynamic environments," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1168–1174, doi: 10.1109/IROS.2018.8593691.

[52] B. Bescós, J. M. Fácil, J. Civera, and J. Neira, "Dynaslam: Tracking, mapping, and inpainting in dynamic scenes," *IEEE Robotics and Automation Letters*, vol. 3, pp. 4076–4083, 2018.

[53] Y. bao Ai, T. Rui, M. Lu, L. Fu, S. Liu, and S. Wang, "Ddl-slam: A robust rgb-d slam in dynamic environments combined with deep learning," *IEEE Access*, vol. 8, pp. 162 335–162 342, 2020.

[54] Y. Fan, Q. Zhang, S. Liu, Y. Tang, X. Jing, J. Yao, and H. Han, "Semantic slam with more accurate point cloud map in dynamic environments," *IEEE Access*, vol. 8, pp. 112 237–112 252, 2020.

[55] Z. Chang, H. Wu, Y. Sun, and C. Li, "Rgb-d visual slam based on yolov4-tiny in indoor dynamic environment," *Micromachines*, vol. 13, no. 2, 2022, doi: 10.3390/mi13020230. [Online]. Available: https://www.mdpi.com/2072-666X/13/2/230

[56] A. Burgdorf, A. Paulus, A. Pomp, and T. Meisen, "Vc-slamdash;a handcrafted data corpus for the construction of semantic models," *Data*, vol. 7, no. 2, 2022, doi: 10.3390/data7020017. [Online]. Available: https://www.mdpi.com/2306-5729/7/2/17

[57] W. Wenxin, G. Liang, G. Hongli, Y. Zhichao, L. Yuekai, and C. Zhiqiang, "Yolo-slam: A semantic slam system towards dynamic environment with geometric constraint," *Springer*, 2022, doi: 10.1007/s00521-021-06764-3.

[58] Y. Fan, Q. Zhang, Y. Tang, S. Liu, and H. Han, "Blitz-slam: A semantic slam in dynamic environments," *Pattern Recognition*, vol. 121, no. 0031-3203, p. 108225, 2022, doi: https://doi.org/10.1016/j.patcog.2021.108225. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320321004064

[59] C. S.-L. Li G.-H., "Visual slam in dynamic scenes based on object tracking and static points detection," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 104, no. 2, 2022, doi: h10.1007/s10846-021-01563-3.

[60] S. Yang, Y. Huang, and S. Scherer, "Semantic 3d occupancy mapping through efficient high order crfs," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 590–597, doi: 10.1109/IROS.2017.8202212.

[61] Y. Nakajima, K. Tateno, F. Tombari, and H. Saito, "Fast and accurate semantic mapping through geometric-based incremental segmentation," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 385–392, doi: 10.1109/IROS.2018.8593993.

[62] M. Runz, M. Buffier, and L. Agapito, "Maskfusion: Real-time recognition, tracking and reconstruction of multiple moving objects," in *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2018, pp. 10–20, doi: 10.1109/ISMAR.2018.00024.

[63] L. Cui and C. Ma, "Sof-slam: A semantic visual slam for dynamic environments," *IEEE Access*, vol. 7, pp. 166 528–166 539, 2019.

[64] A. Iqbal and N. R. Gans, "Data association and localization of classified objects in visual slam," *Journal of Intelligent and Robotic Systems*, pp. 1–18, 2020.

[65] J. Zhang, M. Gui, Q. Wang, R. Liu, J. Xu, and S. Chen, "Hierarchical topic model based object association for semantic slam," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, pp. 3052–3062, 2019.

Page 6

[66] P. Guan, Z. Cao, E. Chen, S. Liang, M. Tan, and J. Yu, "A real-time semantic visual slam approach with points and objects," *International Journal of Advanced Robotic Systems*, vol. 17, 2020.

[67] Z. Zhang, J. Zhang, and Q. Tang, "Mask r-cnn based semantic rgb-d slam for dynamic scenes," *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pp. 1151–1156, 2019.

[68] L. Nicholson, M. Milford, and N. Sünderhauf, "Quadricslam: Dual quadrics from object detections as landmarks in object-oriented slam," *IEEE Robotics and Automation Letters*, vol. 4, no. 1, pp. 1–8, 2019, doi: 10.1109/LRA.2018.2866205.

[69] S. Cascianelli, G. Costante, E. Bellocchio, P. Valigi, M. L. Fravolini, and T. A. Ciarfuglia, "Robust visual semi-semantic loop closure detection by a covisibility graph and cnn features," *Robotics Auton. Syst.*, vol. 92, pp. 53–65, 2017.

[70] V. Murali, H.-P. Chiu, S. Samarasekera, and R. T. Kumar, "Utilizing semantic visual landmarks for precise vehicle navigation," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 1–8, doi: 10.1109/ITSC.2017.8317859.

[71] U.-H. Kim, S.-H. Kim, and J.-H. Kim, "Simvodis++: Neural semantic visual odometry in dynamic environments," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4244–4251, 2022, doi: 10.1109/LRA.2022.3150854.

[72] C. Song, B. Zeng, T. Su, K. Zhang, and J. Cheng, "Data association and loop closure in semantic dynamic slam using the table retrieval method," *Springer*, 2022, doi: 10.1007/s10489-021-03091-x.

[73] A. Rosinol, M. Abate, Y. Chang, and L. Carlone, "Kimera: an open-source library for real-time metric-semantic localization and mapping," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 1689–1696, doi: 10.1109/ICRA40945.2020.9196885.

[74] R. F. Salas-Moreno, R. A. Newcombe, H. Strasdat, P. H. Kelly, and A. J. Davison, "Slam++: Simultaneous localisation and mapping at the level of objects," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1352–1359, doi: 10.1109/CVPR.2013.178.

[75] H. Sang, R. Jiang, Z. Wang, Y. Zhou, and B. He, "A novel neural multi-store memory network for autonomous visual navigation in unknown environment," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2039–2046, 2022, doi: 10.1109/LRA.2022.3140795.

[76] J.-R. Ruiz-Sarmiento, C. Galindo, and J. González, "Exploiting semantic knowledge for robot object recognition," *Knowl. Based Syst.*, vol. 86, pp. 131–142, 2015.

[77] A. Haidu, D. Beßler, A. K. Bozcuoğlu, and M. Beetz, "Knowrobsim — game engine-enabled knowledge processing towards cognition-enabled robot control," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 4491–4498, doi: 10.1109/IROS.2018.8593935.

[78] Y. G. Rocha, S.-H. Joo, E.-J. Kim, and T.-Y. Kuc, "Automatic generation of a simulated robot from an ontology-based semantic description," in *2019 19th International Conference on Control, Automation and Systems (ICCAS)*, 2019, pp. 1340–1343, doi: 10.23919/ICCAS47443.2019.8971767.

[79] Webots, "http://www.cyberbotics.com," open-source Mobile Robot Simulation Software. [Online]. Available: http://www.cyberbotics.com

[80] N. P. Koenig and A. Howard, "Design and use paradigms for gazebo, an open-source multi-robot simulator," *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, vol. 3, pp. 2149–2154 vol.3, 2004.

[81] S. L. Bowman, N. Atanasov, K. Daniilidis, and G. J. Pappas, "Probabilistic data association for semantic slam," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 1722–1729.

[82] Z. Qian, J. Fu, and J. Xiao, "Towards accurate loop closure detection in semantic slam with 3d semantic covisibility graphs," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2455–2462, 2022, doi: 10.1109/LRA.2022.3145066.

[83] N. Dong, M. Qin, J. Chang, C. Wu, W. Ip, and K. Yung, "Weighted triplet loss based on deep neural networks for loop closure detection in vslam," *Computer Communications*, vol. 186, pp. 153–165, 2022, doi: https://doi.org/10.1016/j.comcom.2022.01.013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366422000196

[84] C. Brand, M. J. Schuster, H. Hirschmüller, and M. Suppa, "Submap matching for stereo-vision based indoor/outdoor slam," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 5670–5677.

[85] P. H. Truong, S. You, and S. Ji, "Object detection-based semantic map building for a semantic visual slam system," in *2020 20th International Conference on Control, Automation and Systems (ICCAS)*, 2020, pp. 1198–1201, doi: 10.23919/ICCAS50221.2020.9268441.

[86] J. Yu, Z. Xiang, and J. Su, "Hierarchical multi-level information fusion for robust and consistent visual slam," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 250–259, 2022, doi: 10.1109/TVT.2021.3127554.

[87] A. Nüchter and J. Hertzberg, "Towards semantic maps for mobile robots," *Robotics Auton. Syst.*, vol. 56, pp. 915–926, 2008.

[88] L. Xiao, J. Wang, X. Qiu, Z. Rong, and X. Zou, "Dynamic-slam: Semantic monocular visual localization and mapping based on deep learning in dynamic environment," *Robotics Auton. Syst.*, vol. 117, pp. 1–16, 2019.

[89] M. Grinvald, F. Furrer, T. Novkovic, J. J. Chung, C. Cadena, R. Siegwart, and J. Nieto, "Volumetric instance-aware semantic mapping and 3d object discovery," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 3037–3044, 2019, doi: 10.1109/LRA.2019.2923960.

[90] B. Sean, A. Nikolay, K. Daniilidi, and P. George, "Probabilistic data association for semantic slam," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 1722–1729, doi: 10.1109/ICRA.2017.7989203.

[91] K. Doherty, D. Fourie, and J. Leonard, "Multimodal semantic slam with probabilistic data association," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 2419–2425, doi: 10.1109/ICRA.2019.8794244.

[92] S. Yang and S. Scherer, "Cubeslam: Monocular 3d object detection and slam without prior models," *arXiv preprint*, 2019, doi: arXiv:1806.00557v2.

[93] H. Chen, G. Zhang, and Y. Ye, "Semantic loop closure detection with instance-level inconsistency removal in dynamic industrial scenes," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2030–2040, 2021, doi: 10.1109/TII.2020.3010580.

[94] Y. Wang, Y. Qiu, P. Cheng, and X. Duan, "Robust loop closure detection integrating visual-spatial-semantic information via topological graphs and cnn features," *Remote. Sens.*, vol. 12, p. 3890, 2020.

[95] Z. Liao, Y. Hu, J. Zhang, X. Qi, X. Zhang, and W. Wang, "So-slam: Semantic object slam with scale proportional and symmetrical texture constraints," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4008–4015, 2022, doi: 10.1109/LRA.2022.3148465.

[96] D. D. Zhiwei Xing, Xiaorui Zhu, "De-slam: Slam for highly dynamic environment," *Journal of Field Robotics*, 2022, doi: https://doi-org.wwwproxy1.library.unsw.edu.au/10.1002/rob.22062.

[97] V. Vasilopoulos, G. Pavlakos, K. Schmeckpeper, K. Daniilidis, and D. E. Koditschek, "Reactive navigation in partially familiar planar environments using semantic perceptual feedback," *The International Journal of Robotics Research*, vol. 41, no. 1, pp. 85–126, 2022, doi: 10.1177/02783649211048931. [Online]. Available: https://doi.org/10.1177/02783649211048931

[98] M. Tenorth and M. Beetz, "Knowrob: A knowledge processing infrastructure for cognition-enabled robots," *The International Journal of Robotics Research*, vol. 32, pp. 566 – 590, 2013.

[99] M. Tenorth, L. Kunze, D. Jain, and M. Beetz, "Knowrob-map - knowledge-linked semantic object maps," in *2010 10th IEEE-RAS International Conference on Humanoid Robots*, 2010, pp. 430–435, doi: 10.1109/ICHR.2010.5686350.

[100] C. I. Schlenoff and E. Messina, "A robot ontology for urban search and rescue," in *KRAS '05*, 2005.

[101] Ó. M. Mozos, P. Jensfelt, H. Zender, G.-J. M. Kruijff, and W. Burgard, "From labels to semantics: an integrated system for conceptual spatial representations of indoor environments for mobile robots," 2007.

[102] I. H. Suh, G. H. Lim, W. Hwang, H. Suh, J.-H. Choi, and Y. Park, "Ontology-based multi-layered robot knowledge framework (omrkf) for robot intelligence," *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 429–436, 2007.

[103] M. A. Eid, R. Liscano, and A. E. Saddik, "A universal ontology for sensor networks data," *2007 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, pp. 59–62, 2007.

[104] L. Belouaer, M. Bouzid, and A.-I. Mouaddib, "Ontology based spatial planning for human-robot interaction," *2010 17th International Symposium on Temporal Representation and Reasoning*, pp. 103–110, 2010.

[105] M. Cashmore, M. Fox, D. Long, D. Magazzeni, B. Ridder, A. Carrera, N. Palomeras, N. Hurtós, and M. Carreras, "Rosplan: Planning in the robot operating system," in *ICAPS*, 2015.

[106] E. P. e Silva, J. L. Carbonera, S. R. Fiorini, V. A. M. Jorge, M. Abel, R. Madhavan, A. Locoro, P.-J. S. Gonçalves, M. E. Barreto, M. K. Habib, A. Chibani, S. Gérard, Y. Amirat, and C. I. Schlenoff, "Towards a core ontology for robotics and automation," *Robotics Auton. Syst.*, vol. 61, pp. 1193–1204, 2013.

[107] D. S. Chang, G. H. Cho, and Y. S. Choi, "Ontology-based knowledge model for human-robot interactive services," *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2020.

[108] G. H. Lim, I. H. Suh, and H. Suh, "Ontology-based unified robot knowledge for service robots in indoor environments," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 41, pp. 492–509, 2011.

[109] S. Dhouib, N. du Lac, J.-L. Farges, S. Gérard, M. Hemaissia-Jeannin, J. Lahera-Perez, S. Millet, B. Patin, and S. Stinckwich, "Control architecture concepts and properties of an ontology devoted to exchanges in mobile robotics," 2011.

[110] A. Pronobis and P. Jensfelt, "Multi-modal semantic mapping," 2011.

[111] T. Wang and Q. J. Chen, "Object semantic map representation for indoor mobile robots," *Proceedings 2011 International Conference on System Science and Engineering*, pp. 309–313, 2011.

[112] S. von Riegen, L. Hotz, and P. Rost, "Combining qualitative spatial reasoning and ontological reasoning for supporting robot tasks," in *KEOD*, 2012.

[113] L. Paull, G. Séverac, G. V. Raffo, J. M. Angel, H. Boley, P. J. Durst, W. Gray, M. K. Habib, B. Nguyen, S. V. Ragavan, G. SajadSaeedi, R. Sanz, M. L. Seto, A. Stefanovski, M. Trentini, and H. Li, "Towards an ontology for autonomous robots," *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1359–1364, 2012.

[114] R. Li, L. Wei, D. Gu, H. Hu, and K. D. Mcdonald-Maier, "Multi-layered map based navigation and interaction for an intelligent wheelchair," *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 115–120, 2013.

[115] J. L. Carbonera, S. R. Fiorini, E. P. e Silva, V. A. M. Jorge, M. Abel, R. Madhavan, A. Locoro, P.-J. S. Gonçalves, T. Haidegger, M. E. Barreto, and C. I. Schlenoff, "Defining positioning in a core ontology for robotics," *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1867–1872, 2013.

[116] H. Wu, G. Tian, Y. Li, F. Zhou, and P. Duan, "Spatial semantic hybrid map building and application of mobile service robot," *Robotics Auton. Syst.*, vol. 62, pp. 923–941, 2014.

[117] L. Riazuelo, M. Tenorth, D. D. Marco, M. Salas, D. Gálvez-López, L. Mösenlechner, L. Kunze, M. Beetz, J. D. Tardós, L. Montano, and J. M. M. Montiel, "Roboearth semantic mapping: A cloud enabled knowledge-based approach," *IEEE Transactions on Automation Science and Engineering*, vol. 12, pp. 432–443, 2015.

[118] G. Burroughes and Y. Gao, "Ontology-based self-reconfiguring guidance, navigation, and control for planetary rovers," *J. Aerosp. Inf. Syst.*, vol. 13, pp. 316–328, 2016.

[119] F. Ramos, A. S. Vázquez, R. Fernández, and A. O. Alarcos, "Ontology based design, control and programming of modular robots," *Integr. Comput. Aided Eng.*, vol. 25, pp. 173–192, 2018.

[120] H. Deeken, T. Wiemann, and J. Hertzberg, "Grounding semantic maps in spatial databases," *Robotics Auton. Syst.*, vol. 105, pp. 146–165, 2018.

[121] X. Sun, Y. Zhang, and J. Chen, "High-level smart decision making of a robot based on ontology in a search and rescue scenario," *Future Internet*, vol. 11, p. 230, 2019.

[122] J. Crespo, J. C. Castillo, Ó. M. Mozos, and R. Barber, "Semantic information for robot navigation: A survey," *Applied Sciences*, vol. 10, p. 497, 2020.

[123] S.-H. Joo, S. Manzoor, Y. G. Rocha, S.-H. Bae, K. hee Lee, T.-Y. Kuc, and M. Kim, "Autonomous navigation framework for intelligent robots based on a semantic environment modeling," *Applied Sciences*, 2020.

[124] S. Karimi, I. Iordanova, and D. St-Onge, "An ontology-based approach to data exchanges for robot navigation on construction sites," *J. Inf. Technol. Constr.*, vol. 26, pp. 546–565, 2021.

[125] S. V. Shchekotov, N. Smirnov, and M. Pashkin, "The ontology driven slam based indoor localization technique," *Journal of Physics: Conference Series*, vol. 1801, 2021.

[126] T. Ben, D. Feras, C. Peter, and W. Gordon, "Robot navigation in unseen spaces using an abstract map," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 4, pp. 791–805, 2021, doi: 10.1109/TCDS.2020.2993855.

[127] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007, doi: 10.1.1.117.471. [Online]. Available: https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf

[128] P. Matthew, M. Joanne, B. Patrick, B. Isabelle, H. Tammy, M. Cynthia, S. Larissa, T. Jennifer, A. Elie, B. Sue, C. Roger, G. Julie, G. Jeremy, H. Asbjorn, L. Manoj, L. Tianjing, L. Elizabeth, M.-W. Evan, M. Steve, M. Luke, S. Lesley, T. James, T. Andrea, W. Vivian, W. Penny, and M. David, "The prisma 2020 statement: an updated guideline for reporting systematic reviews," vol. 10, p. 89, 2021, doi: 10.1186/s13643-021-01626-4,. [Online]. Available: https://doi.org/10.1186/s13643-021-01626-4

[129] A. R. Washburn, *Search and detection*, 1st ed. Operations Research Department, Naval Postgraduate School, 1981.

[130] A. Subasi, "Chapter 4 - feature extraction and dimension reduction," in *Practical Guide for Biomedical Signals Analysis Using Machine Learning Techniques*, A. Subasi, Ed. Academic Press, 2019, pp. 193–275, doi: https://doi.org/10.1016/B978-0-12-817444-9.00004-0. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128174449000040

[131] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10 778–10 787, 2020.

[132] S. Zhang, C. Chi, Y. Yao, Z. Lei, and S. Z. Li, "Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9756–9765, 2020.

[133] Y. Lee and J. Park, "Centermask: Real-time anchor-free instance segmentation," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13 903–13 912, 2020.

[134] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *ArXiv*, vol. abs/2004.10934, 2020.

[135] B. Williams, M. Cummins, J. Neira, P. Newman, I. Reid, and J. Tardós, "A comparison of loop closing techniques in monocular slam," *Robotics and Autonomous Systems*, vol. 57, no. 12, pp. 1188–1197, 2009, inside Data Association. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0921889009000876

[136] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM: a versatile and accurate monocular SLAM system," *CoRR*, vol. abs/1502.00956, 2015. [Online]. Available: http://arxiv.org/abs/1502.00956

[137] R. Mur-Artal and J. D. Tardós, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017, doi: 10.1109/TRO.2017.2705103.

[138] D. Galvez-López and J. D. Tardos, "Bags of binary words for fast place recognition in image sequences," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1188–1197, 2012, doi: 10.1109/TRO.2012.2197158.

[139] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *ECCV*, 2006.

[140] M. Calonder, V. Lepetit, C. Strecha, and P. V. Fua, "Brief: Binary robust independent elementary features," in *ECCV*, 2010.

[141] E. Rublee, V. Rabaud, K. Konolige, and G. R. Bradski, "Orb: An efficient alternative to sift or surf," *2011 International Conference on Computer Vision*, pp. 2564–2571, 2011.

[142] T. Barros, R. Pereira, L. Garrote, C. Premebida, and U. J. Nunes, "Place recognition survey: An update on deep learning approaches," *CoRR*, vol. abs/2106.10458, 2021. [Online]. Available: https://arxiv.org/abs/2106.10458

[143] R. Arandjelović, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "Netvlad: Cnn architecture for weakly supervised place recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1437–1451, 2018, doi: 10.1109/TPAMI.2017.2711011.

[144] S. Garg, N. Suenderhauf, and M. Milford, "Semantic–geometric visual place recognition: a new perspective for reconciling opposing views," *The International Journal of Robotics Research*, vol. 0, no. 0, p. 0278364919839761, 2019, doi: 10.1177/0278364919839761. [Online]. Available: https://doi.org/10.1177/0278364919839761

[145] S. Garg, N. Sünderhauf, and M. Milford, "Lost? appearance-invariant place recognition for opposite viewpoints using visual semantics," *CoRR*, vol. abs/1804.05526, 2018. [Online]. Available: http://arxiv.org/abs/1804.05526

[146] Y. Kong, W. Liu, and Z. Chen, "Robust convnet landmark-based visual place recognition by optimizing landmark matching," *IEEE Access*, vol. 7, pp. 30 754–30 767, 2019, doi: 10.1109/ACCESS.2019.2901984.

[147] G. Singh, M. Wu, S. Lam, and D. Van Minh, "Hierarchical loop closure detection for long-term visual slam with semantic-geometric descriptors," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021, pp. 2909–2916, doi: 10.1109/ITSC48978.2021.9564866.

[148] N. Sünderhauf, S. Shirazi, A. Jacobson, F. Dayoub, E. Pepperell, B. Upcroft, and M. Milford, "Place recognition with convnet landmarks: Viewpoint-robust, condition-robust, training-free," in *Robotics: Science and Systems*, vol. 11, 2015.

[149] B. Mu, S.-Y. Liu, L. Paull, J. Leonard, and J. P. How, "Slam with objects using a nonparametric pose graph," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 4602–4609, doi: 10.1109/IROS.2016.7759677.

[150] J. Dong, E. Nelson, V. Indelman, N. Michael, and F. Dellaert, "Distributed real-time cooperative localization and mapping using an uncertainty-aware expectation maximization approach," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 5807–5814, doi: 10.1109/ICRA.2015.7140012.

[151] S. Yadav, D. Mahato, and N. Linh, *Distributed Artificial Intelligence: A Modern Approach (1st ed.)*. CRC Press, 2020, doi: https://doi.org/10.1201/9781003038467.

[152] L. Jean-Baptiste, *Ontologies with Python*, 1st ed. Apress, 2021.

[153] A. Gomez-Perez, M. Fernandez-Lopez, and O. Corcho, *Ontological engineering*, 1st ed. Springer, 2010.

[154] A. J. Hepworth, D. P. Baxter, and H. A. Abbass, "Onto4mat: A swarm shepherding ontology for generalised multi-agent teaming," 2022, doi: 10.48550/ARXIV.2203.12955. [Online]. Available: https://arxiv.org/abs/2203.12955

[155] A. Hussein, S. Elsawah, and H. A. Abbass, "Trust mediating reliability–reliance relationship in supervisory control of human–swarm interactions," *Human Factors*, vol. 62, no. 8, pp. 1237–1248, 2020, doi: 10.1177/0018720819879273. [Online]. Available: https://doi.org/10.1177/0018720819879273

[156] C. Zhu, F. Chen, U. Ahmed, Z. Shen, and M. Savvides, "Semantic relation reasoning for shot-stable few-shot object detection," *CVPR*, 2021.

[157] Y. Zhang, M. Jiang, and Q. Zhao, "Explicit knowledge incorporation for visual reasoning," *CVPR*, 2021.

[158] L. Riazuelo, M. Tenorth, D. Di Marco, M. Salas, D. Gálvez-López, L. Mösenlechner, L. Kunze, M. Beetz, J. D. Tardós, L. Montano, and J. M. M. Montiel, "Roboearth semantic mapping: A cloud enabled knowledge-based approach," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 2, pp. 432–443, 2015, doi: 10.1109/TASE.2014.2377791.

[159] M. Tenorth, A. Clifford Perzylo, R. Lafrenz, and M. Beetz, "The roboearth language: Representing and exchanging knowledge about actions, objects, and environments," in *2012 IEEE International Conference on Robotics and Automation*, 2012, pp. 1284–1289, doi: 10.1109/ICRA.2012.6224812.

[160] H. Wu, C. Du, Z. Ji, M. Gao, and Z. He, "Sort-ym: An algorithm of multi-object tracking with yolov4-tiny and motion prediction," *Electronics*, 2021.

[161] S. Han and Z. Xi, "Dynamic scene semantics slam based on semantic segmentation," *IEEE Access*, vol. 8, pp. 43 563–43 570, 2020.

[162] L. Zhang, L.-Y. Wei, P. Shen, W. Wei, G. Zhu, and J. Song, "Semantic slam based on object detection and improved octomap," *IEEE Access*, vol. 6, pp. 75 545–75 559, 2018.

[163] P. Li, G. Zhang, J. Zhou, R. Q. Yao, X. Zhang, and J. Zhou, "Study on slam algorithm based on object detection in dynamic scene," *2019 International Conference on Advanced Mechatronic Systems (ICAMechS)*, pp. 363–367, 2019.

[164] T. BRAUNL, *ROBOT ADVENTURES IN PYTHON AND C*, 1st ed. SPRINGER, 2021, doi: https://doi.org/10.1007/978-3-030-38897-3.

[165] H. W. Yu, J. Y. Moon, and B. H. Lee, "A variational observation model of 3d object for probabilistic semantic slam," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 5866–5872.