

Article

Ontology-Based Deep Learning Model for Object Detection and Image Classification in Smart City Concepts

Adekanmi Adeyinka Adegun ^{1,2,*} , Jean Vincent Fonou-Dombeu ^{3,†} , Serestina Viriri ^{2,*}  and John Odindi ^{4,†} 

¹ Department of Computing, School of Arts, Humanities, and Social Sciences, University of Roehampton, London SW15 5PH, UK

² School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban 4000, South Africa

³ School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Pietermaritzburg 3209, South Africa; fonoudombeuj@ukzn.ac.za

⁴ School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg 3209, South Africa; odindij@ukzn.ac.za

* Correspondence: adekanmi.adekun@roehampton.ac.uk (A.A.A.); viriris@ukzn.ac.za (S.V.)

† These authors contributed equally to this work.

Abstract: Object detection in remotely sensed (RS) satellite imagery has gained significance in smart city concepts, which include urban planning, disaster management, and environmental monitoring. Deep learning techniques have shown promising outcomes in object detection and scene classification from RS satellite images, surpassing traditional methods that are reliant on hand-crafted features. However, these techniques lack the ability to provide in-depth comprehension of RS images and enhanced interpretation for analyzing intricate urban objects with functional structures and environmental contexts. To address this limitation, this study proposes a framework that integrates a deep learning-based object detection algorithm with ontology models for effective knowledge representation and analysis. The framework can automatically and accurately detect objects and classify scenes in remotely sensed satellite images and also perform semantic description and analysis of the classified scenes. The framework combines a knowledge-guided ontology reasoning module into a YOLOv8 objects detection model. This study demonstrates that the proposed framework can detect objects in varying environmental contexts captured using a remote sensing satellite device and incorporate efficient knowledge representation and inferences with a less-complex ontology model.

Keywords: smart city; urban planning; image analysis; remote sensing; satellite images; object detection; deep learning; YOLO; knowledge representation; ontology



Citation: Adegun, A.A.; Fonou-Dombeu, J.V.; Viriri, S.; Odindi, J. Ontology Based Deep Learning Model for Object Detection and Image Classification in Smart City Concept. *Smart Cities* **2024**, *7*, 2182–2207. <https://doi.org/10.3390/smartcities7040086>

Academic Editor: Pierluigi Siano

Received: 1 June 2024

Revised: 23 July 2024

Accepted: 25 July 2024

Published: 2 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Remote sensing (RS) image analysis plays a crucial role in various applications that include land cover mapping, urban planning, and environmental monitoring [1]. Remote sensing image scene detection and classification involve identifying and categorizing different land cover types and urban features from satellite images [2]. In urban landscapes, these tasks are crucial for land management, resource monitoring, and sustainable development. Recently, various methods have been used for object detection and scene classification from remotely sensed satellite images. However, traditional methods for scene detection and classification, which rely on manual feature extraction and handcrafted algorithms, are often unable to accurately capture the complex spatial and spectral patterns in remotely sensed data due to their lack of accuracy and scalability [3]. Whereas state-of-the-art deep learning techniques have lately shown significant potential and improvements in automatically learning relevant features for efficient scene detection and classification, they are limited in accurate recognition, interpretability, and semantic analysis of the objects and scenes detected [4,5]. This is due to the composition of RS images, which include complex

urban objects that vary in geometrical shapes, functional structures, environmental contexts, and semantic heterogeneity.

Deep learning models, such as convolutional neural networks (CNNs), have demonstrated exceptional performance in various computer vision tasks, including image classification and object detection [6]. In remote sensing image analysis, deep learning models have been employed to learn complex spatial and spectral patterns from large-scale datasets [7]. These models automatically extract hierarchical representations, allowing for more effective scene detection and classification. The methods have been used for automatic analysis of complex images to identify and classify disaster management, planning, and mitigation for their relevant object contents [8]. For example, a CNN-LSTM model combination was proposed by Husni et al. [9] for identifying littering behavior in smart environments. A tiny garden and a river location were used to evaluate the proposed method. A model was also used to evaluate the flood risk of some bounded metropolitan areas using geospatial data from publicly accessible databases [10]. Vehicle detection in a variety of environments was assessed by Shokri et al. [11] using object detection-based deep learning methods. Nadeem et al. [12] developed FlameNet, a CNN-based system, for fire detection in a smart city setting. The network was built with the ability to detect fires and send an alarm to the relevant agency. An explainable AI was utilized by Thakker et al. [13] as a hybrid classifier for object detection and categorization of images with the presence of flooding by fusing deep learning and semantic web technologies. The system offered flexibility in classifying images and detecting objects using their coverage connections and expert knowledge. Cong et al. [14] performed a precision analysis of geotechnical investigations and urban planning by combining smart sensing technology with a predictive analytics method utilizing Kriging and ensemble learning. Lastly, Chen et al. [15] designed a decoder network for the semantic segmentation of the Adverse Conditions Dataset collected with Correspondences (ACDC) using images obtained in various smart city environments at different times and locations.

The development of knowledge-driven approaches that incorporate remote sensing image analysis with expert knowledge from environmental scientists for effective interpretation of the images has been identified as an important direction in data-driven research [16]. In recent years, the integration of ontology-based techniques with machine learning algorithms has shown promising results in semantic understanding of scene detection and object classification tasks [17,18]. Ontology provides a formal representation of knowledge in a domain, enabling semantic understanding and reasoning. Applying ontology in remote sensing image analysis is important in capturing essential knowledge such as spatial relationships and semantic hierarchies in land cover classes [19]. This enhances automatic semantic analysis to aid in more interpretable scene detection and classification. Combining deep learning's ability to learn complex patterns and ontologies' contextual understanding will enhance the effectiveness of scene detection and classification tasks and ultimately achieve robust land cover mapping, urban planning, environmental monitoring, and contribute to sustainable development and resource management [20,21]. Hence, ontology-based deep learning offers a promising approach for remote sensing image scene detection and classification. Integrating ontology with deep learning models by effectively incorporating semantic knowledge into the image analysis process has led to improved accuracy and interpretability [22].

Despite the potential benefits of ontology-based deep learning, several challenges exist. Constructing and maintaining ontology for remote sensing image analysis can be complex and time-consuming [23]. Additionally, the semantic gap between the ontology and the deep learning models poses a challenge in their effective integration. Bridging this gap requires careful ontology design and alignment with the deep learning architecture [23]. Nevertheless, the integration of ontology-based techniques with deep learning opens up exciting opportunities. This study aims to achieve fine-grained objects detection and scene classification, improved interpretability, and ultimately contribute to more accurate land cover mapping, urban planning, and environmental monitoring for disaster planning. This

study presents an ontology-based deep learning framework that combines a deep learning algorithm and ontology for automatic detection and classification of objects in remotely sensed satellite images and performs knowledge extraction and modeling of the scene classification outcomes. By combining the semantic knowledge from ontology with the representation and prediction capabilities of deep learning models, more interpretable scene detection and classification can be achieved. The proposed model employs an efficient computer vision algorithm for object detection from remotely sensed satellite images captured from the Kwazulu-Natal province, South Africa. This is combined with an ontology model for semantic reasoning of detected objects in the scenes. The ontology provides a framework for capturing domain-specific knowledge, improving the understanding of remote sensing images, and enhancing the interpretability of deep learning models.

The main contributions of this study include:

1. Modeling a lightweight YOLOv8-based algorithm for object detection and scene classification in remotely sensed satellite images.
2. Proposing a mathematical and ontological model for establishing knowledge taxonomy, knowledge extraction, and inferences.
3. Creating SPARQL database for knowledge inferences and validation of the ontological reasoning. The SPARQL database is made publicly available by the authors, as shared in the Data Availability Statement, at: <https://data.mendeley.com/datasets/s5v4zz7yj5/1> (accessed on 1 June 2024).

The remaining part of this work is organized as follows: Section 2 discusses a review of related works. Section 3 describes the research methodology and ontology modeling. Section 4 discusses the results, summary and future works. This paper is concluded in Section 5.

2. Review of Related Works

This section highlights various approaches used in integrating ontology-based frameworks with image analysis for knowledge representation and inferences. This has gained attention in recent years, with machine learning models incorporating ontological modeling in the image analysis process. An image classification process was modeled using ontology to represent decision tree-based classifiers and rule-based expert systems to provide a data-sharing mechanism with applications working on oceanic images [24]. Alirezaie et al. [25] developed an ontology framework, SemCityMap, for satellite image classification. The framework incorporated semantic information about mapping locations and paths for the provision of a knowledge representation and reasoning methods to achieve high-level querying. Also, image classification based on ontology and a hierarchical max-pooling (HMAX) model that utilized merged classifiers was proposed by Filali et al. [26]. Ontological relationships between image categories that were in line with training visual feature classifiers were derived by merging the outputs of hypernym–hyponym classifiers for better discrimination of the classes detected. Wang et al. [27] also developed an ontology-based framework for integrating remotely sensed imagery, image products, and in situ observations. The system combined remote sensing imagery with semantic queries based on a description logics (DL) query and SPARQL.

A deep learning algorithm integrated with ontology models was employed by Fang et al. [28] to develop a knowledge graph that can recognize falling from height (FFH) hazards. The ontological model employed knowledge extraction and knowledge inference based on image analysis and classification for hazard prediction. Miranda et al. [29] also proposed an ontology-based deep learning framework for land cover classification. They used a semantic reasoning approach in a medium-resolution optical imagery document containing features such as the normalized difference vegetation index, brightness, gray level co-occurrence matrix homogeneity, and rectangular fit from Indonesia National standard RSNI-1 Land Cover satellite imagery. They demonstrated that incorporating the semantic information from the ontology improved the classification accuracy of Sentinel-2 satellite imagery. Xie et al. [22] presented an ontology-based methodology framework for

enabling object recognition using rules extracted from high-level semantics. The framework was able to semantically organize the descriptions and definitions of objects using RDF-triple semantic rules from the developed domain ontology. Low-level data features defined from optical satellite and LiDAR images were mapped to the decomposed parts of the RDF-triple rules, and a probabilistic belief network (PBN) was used to represent the relationships between low-level data features and high-level semantics.

Sambandam et al. [30] also used a semantic web technology to establish the spatial ontology for risk knowledge in spatial datasets. The model was integrated with deep attention-based bidirectional search and rescue long short-term memory for efficient image analysis. The system was experimented on using the University of California Merced (UCM) dataset, with an overall accuracy (OA) of 92.3%. Also, an ontology-driven hierarchical sparse representation for hierarchical learning for large-scale image classification was developed [31]. The system used WordNet to construct semantic ontology in the form of a visual ontology tree based on deep features extracted by Inception V3. An algorithm based on split Bregman iteration was developed to learn hierarchical sparse representation and was evaluated using three benchmark datasets: ILSVRC2010, SUN397, and Caltech256.

Benkirane et al. [32] proposed an ontology model in a deep learning context for representing urban environments using a structured set of concepts connected by semantic relationships. The ontology model was used to extract monocular cue information from images, which was sent, together with the images, to a deep neural network model for knowledge inferential analysis. The system was experimented on using benchmark datasets: KITTI, CityScapes, and AppolloScape. A deep learning-based ontology model was used to generate a set of features trained using ontology for the image classification process [33]. In the system, an ontological bagging algorithm was integrated with an ensemble technique of convolutional neural network (CNN) models to improve forest image classification accuracy. The ensemble technique, which was composed of ResNet50, VGG16, and Xception, achieved a 96% accuracy and 53.2% root mean square error when experimenting on the forest dataset. An ontology-guided deep learning approach for urban land use classification was proposed by Li et al. [18] using a combination of a collaboratively boosting framework (CBF) with a data-driven deep learning module and a knowledge-guided ontology reasoning module. The ontology reasoning module was composed of both intra- and extra-taxonomy reasoning models for correcting misclassifications and generating inferred channels to improve the discriminative performance of DSSN in the original remote sensing (RS) image space. The system was experimented on using two publicly open RS datasets: UCM and ISPRS Potsdam. Gupta et al. [34] developed an algorithm that combined CNN and ontology for inferring abstract patterns in Indian monument images. A transfer learning-based approach was used, in which domain knowledge was transferred to a CNN during training via top-down transfer and inference was made using CNN prediction and an ontology tree via bottom-up transfer. Kim et al. [35] developed a scene graph generation method based on the RDF model to establish semantic relations in images. Deep learning models were used to generate scene graphs expressed in a controlled vocabulary, improving the understanding of relations between image objects.

A lightweight CNN model based on a deep neural network was employed in an object detection process for surface scratch detection [36]. Fundamental semantic sensor network (SSN) ontologies for a fire prediction and management system were proposed by Chandra et al. [37]. Information on several meteorological conditions, such as temperature, relative humidity, and wind speed, was gathered using the semantic sensor networks. In order to calculate fire weather indices, the system used ontology rules, which SPARQL then translated into a resource description framework (RDF). Li et al. [38] proposed a framework that integrates computer vision, ontology, and natural language processing for enhancing systematic safety management for hazard avoidance and elimination. Patel et al. [39] developed a system for locating concealed, abandoned bags in public areas. The system identified and predicted various interactions between the items in images using computer vision-based visual connection identification. Salient information in video

footage was extracted and represented as a knowledge graph using the suggested ontology-based method. Using SPARQL queries, the unexpected events were found based on the knowledge retrieved. The ABODA Dataset, AVSS 2007, PETS 2006, and PETS 2007 were used for testing the propose system. A strategy to boost confidence in machine learning models used in safety-critical domains was proposed by Lynn et al. [40]. The resilience and completeness of the model's training dataset were guaranteed by the system's design. The proposed approach used a combination of domain ontology and characteristic ontologies for image quality to validate the training dataset in terms of image quality and domain completeness. A scene graph engineering and reasoning method based on ontology was presented to explain the structural links derived from retrieved objects by Raj et al. [41]. The ontological model in the proposed system produced associated entities and relationships from objects detected using YOLO. The system used the semantic web rule language (SWRL) to find the image sequence for the structural link.

Lastly, it can be concluded from this review that limited works have been performed regarding the integration of the ontology model with state-of-the-art deep learning methods for the analysis of complex urban objects with functional structures and environmental contexts. Traditional machine learning image analysis algorithms have shown some weaknesses in the detection and classification of objects in complex images [42,43] such as remotely sensed satellite images. In fact, the traditional feature extraction methods in complex images do not capture the semantic relationships between the objects contained in the images, making it difficult to represent semantic knowledge in images with ontology. The application of deep learning-based approaches integrated with ontology modeling has been limited to image classification tasks. To the best of our knowledge, ontology modeling has been rarely applied to semantic analysis of deep learning-based object detection model outputs from remote-sensing satellite images for the prevention of disaster. This study therefore explores the possibility of utilizing ontology for the semantic analysis of deep learning-based object detection outputs. Some ontology frameworks and their limitations are summarized in Table 1.

Table 1. Ontology frameworks with various image datasets and their limitations.

Methods	Datasets	Limitations
Ontology with decision tree-based classifiers and rule-based expert systems [24]	Oceanic images	The framework is limited to traditional machine learning and image classification
Qualitative spatial reasoning framework [25]	Satellite images	This framework does not involve machine learning algorithms
Mask R-CNN with ontology models [28]	Satellite images	Object detection accuracy is low, with high error rates
Ontology and HMAX model [26]	ImageNet	Framework limited to traditional machine learning and image classification
Description Logics (DL) Query and SPARQL [27]	Remote sensing imagery	Framework employs only Description Logics rules
Deep learning, NDVI, brightness, GLCM [29]	Indonesia National Standard RSNI-1 land cover satellite imagery	The study performs land use classification with ontology rule
RDF-triple rules and probabilistic belief network (PBN) [22]	Optical satellite and LiDAR images	The work is limited to traditional machine learning and image classification
Bidirectional search and rescue LSTM [30]	UCM dataset	Combines NLP with a search algorithm; requires complex algorithm
Inception V3 and split Bregman iteration algorithm [31]	ILSVRC2010, SUN397, and Caltech256	Requires large computational resources, performs only image classification
Deep neural network model [32]	KITTI, CityScapes, and AppolloScape	Requires large computational resources, performs only image classification
Ensemble technique of ResNet50, VGG16, and Xception [33]	Forest dataset	Requires large computational resources, performs only image classification
Collaboratively boosting framework (CBF) with data-driven deep learning [18]	UCM and ISPRS Potsdam	Requires large computational resources, performs image classification
CNN and ontology [34]	Indian monument images	performs image classification
RDF, CNN, and RNN deep learning models [35]	Digital images	Requires larger computational resources, lower detection speed and accuracy

3. Materials and Methods

3.1. Methods Overview

The proposed methodology employs the integration of deep learning-based objects detection techniques with ontology, which represents domain knowledge of the predicted output from the object detection model. This enhances the system by providing contextual information and domain knowledge via ontology, leading to more accurate and interpretable scene detection and classification for efficient remote sensing image scene detection and classification. This approach is beneficial for the semantic understanding and the categorization of scenes captured in remote sensing imagery. The system gives further details such as spatial relationships and semantic hierarchies in detected objects classes. Figure 1 illustrates the data flow processes of the proposed system. Training datasets are first collected and annotated appropriately for machine learning modeling. The well-labeled datasets are used to train the proposed deep learning models for object detection and scene classification. Figure 2 shows that the input images are taken through the process of feature extraction, feature fusing, and feature detection in the object detection model. The outcomes from this unit are sent into the ontology development unit, where further semantic analysis is carried out. The results from the ontology units are analyzed and presented as the final outputs. The processes involved in our methodology are highlighted below:

- Dataset creation and preprocessing
- Deep learning models for object detection and scene classification
- Integrating deep learning model with the ontology model
- Ontology modeling

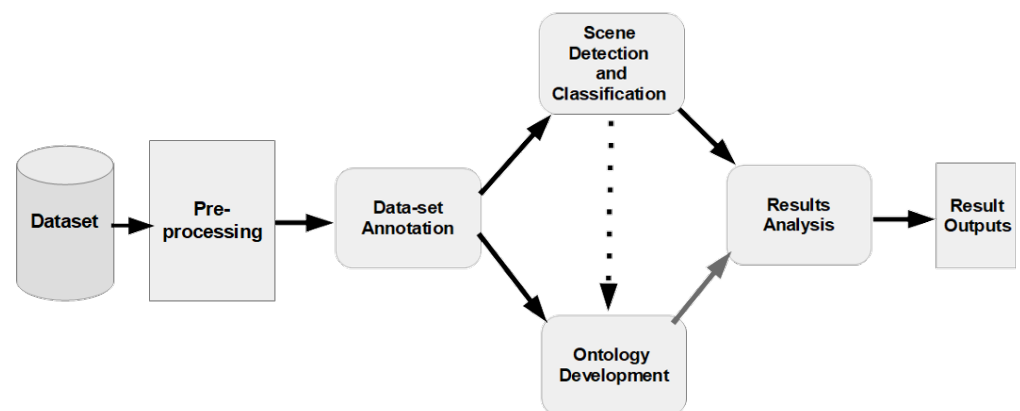


Figure 1. Data flow layout for the proposed methodology.

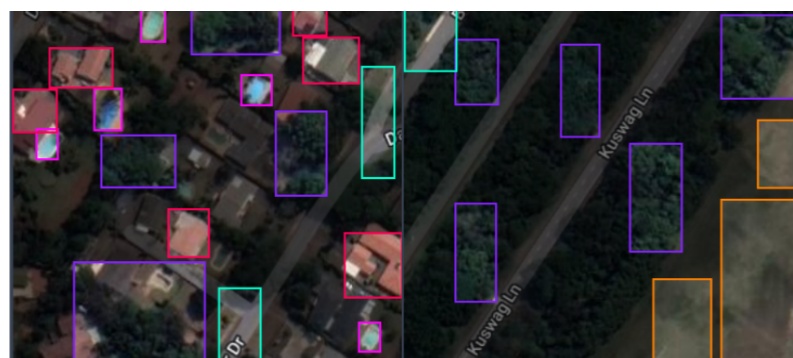


Figure 2. Training sample showing objects in remote-sensed satellite image.

3.2. Dataset Creation

The dataset creation process is the first step of the system development pipeline, as presented in Figure 1. It involves processes such as data collection and preprocessing. In this study, the steps involved are highlighted below:

- **Data collection:** This is the process of gathering and collecting domain-specific high-resolution images from remotely sensed satellite imagery into a training dataset. The images are acquired for the dataset creation process using Google Earth Engine [44,45]. The resolution of the preprocessed images is set to 640×640 pixels. Sentinel-2 MSI imagery from the Kwazulu-Natal region of South Africa's Southern Durban metropolis (coordinates: latitude -29.8579 and longitude 31.0292) is acquired using the engine. To organize and annotate the images for the creation of datasets, Roboflow [46] is utilized as an end-to-end computer vision platform. Roboflow presents a step-by-step approach to achieve a well-labeled dataset for a multi-class classification tasks in which the five objects required to be detected by the machine learning model are first highlighted and initialized. Ninety-two (92) satellite images are then annotated using the multi-class classification procedure in Roboflow, making up the dataset. Five objects are selected from the images, including residences, roads, shorelines, swimming pools, and vegetation. The dataset is also divided into three sets: testing, validation, and training. The training set comprises one annotation file and sixty-one photos. The validation set contains twenty-one images and one annotation file, and the testing set contained ten images. The dataset is also subjected to augmentation processes, increasing them by 100.
- **Image preprocessing:** The collected images are preprocessed by applying necessary corrections, such as radiometric and geometric corrections, to ensure accurate and consistent inputs for the deep learning model. The images are resized at 640×640 dimensions each and annotated in multi-class format to identify five objects: residences, roads, shorelines, swimming pools, and vegetation from each image. A sample image showing some of the objects is presented in Figure 2.
- **Data modeling:** This involves defining and modeling the dataset using ontology with domain-specific knowledge, which focuses on land cover types such as vegetation, water bodies, urban areas, and other scene categories relevant to the region.
- **Data augmentation:** The dataset is augmented to increase the volume for effective model training. The tasks involved in the augmentation process include the following:
 - Auto-orientation of the images in the dataset
 - Geometric augmentations and transformations that change the spatial orientation of the images
 - Rotating the image vertically to reverse it and flipping it horizontally to generate an identical copy
 - Images are rotated by 90, 180, or 270 degrees to represent various viewing angles of an object.

These procedures are performed several times to increase the number of training images by 100. These augmentation techniques are applied to the dataset to improve the model's performance, detection accuracy, and robustness. The methods achieve diversity of the dataset and strengthen the model's resistance to variations in perspective or orientation. The dataset, after augmentation, is further categorized into three subsets: training, validation, and testing. The whole training set contains 6100 images and one annotation file, the validation set contains 2100 images and one annotation file, and the testing set contains 1000 images.

3.3. Deep Learning Model for Object Detection and Scene Classification

The proposed methodology presented in Figure 3 combines the power of deep learning-based object detection methods with the semantic knowledge captured in the ontology for comprehensive understanding and categorizing of scenes captured in remotely sensed

imagery. This approach can enhance scene detection and classification in remotely sensed imagery, specifically tailored to the unique land cover characteristics and requirements.

A well-labeled and annotated training dataset with the corresponding scene categories sourced from the remote satellite dataset is utilized to train the deep learning-based object detection model. In this study, a robust deep learning architecture suitable for scene detection and classification, a lightweight YOLOv8-based deep learning model [47], is employed for object detection from the remotely sensed satellite images collected. As presented in Figure 3, the object detection model is categorized into three (3) units: backbone, neck, and head. The backbone unit is used for feature extraction, the neck unit is used for feature concatenation, and the head unit is used for detection and prediction. The unit compositions are described below:

- Backbone: a set of convolution layers, coarse to fine (C2f) modules, and spatial pyramid pooling faster (SPPF) modules is used as the backbone for feature extraction.
- Neck: the C2f module is used in the neck region to replace the traditional CSP and C3 modules in earlier models to achieve state-of-the-art (SOTA) performance for the YOLOv8-based deep learning model [48].
- Heads: The YOLOv8 detection heads involve anchor-free detection and comprise detection modules and a prediction layer. The two decoupled segmentation heads in the head structure follow the C2f component [48].

The architecture effectively replaces the CSP modules in the neck and backbone areas with a novel C2f module concentrating on specific features exclusively, resulting in a lightweight framework with a reduced number of parameters and overall size of the tensors.

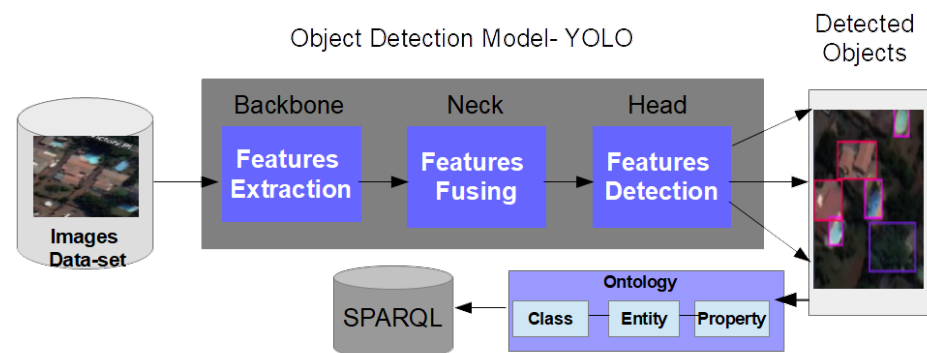


Figure 3. Diagram showing the framework for the proposed ontology-based deep learning system.

3.4. Model Integration

As illustrated in Figure 3, input images are first sent into the deep learning YOLO model for object detection and extraction. The detected and extracted objects are then sent into the SPARQL [49] database for further processing of knowledge extraction. A SPARQL database [50] is created to accept as input the objects extracted from the images that were already analyzed in the deep learning model. SPARQL queries use ontological rules for semantically analyzing data input to establish the hierarchical order of the objects and knowledge expression discovered using the ontology model. The full integration is illustrated in Figures 1 and 3.

3.5. Ontology Modeling and Developments

Ontology modeling and the development of the knowledge extracted from the output of the deep learning-based object detection model is executed in three stages. As presented in Figure 4, the three main processes involve ontology modeling, knowledge extraction, and knowledge inference. The first process involves the creation of a knowledge base

and establishing the ontology taxonomy. The taxonomies identified are represented using mathematical models. The second stage involves the definition of classes and their associated objects, attributes, and events. In the last stage, the extracted knowledge is stored in the form of graph data for the purpose of knowledge inference. The ontology model is designed using RDF and the OWL [49] language and queried using the SPARQL protocol and RDF query language (SPARQL) [49].

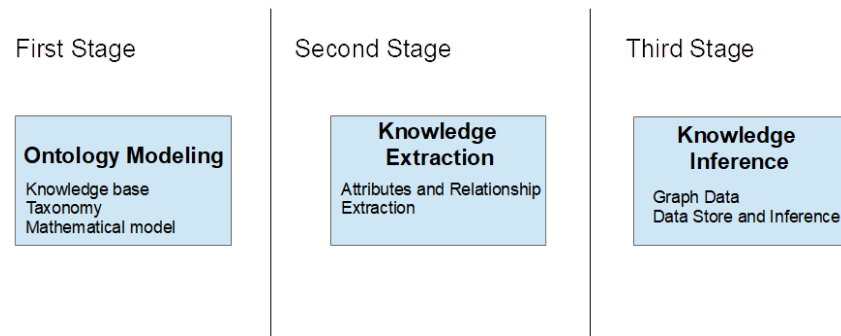


Figure 4. Diagram showing the processes in ontology modeling.

The major stage in the ontology modeling and development process involves three processes. They are highlighted and discussed below:

- Knowledge base creation
- Establishing knowledge taxonomy
- Ontology concept description using mathematical modeling

3.5.1. Knowledge Base Creation

The information required for the creation of the knowledge base revolves around scenes classification and object detection. The main class identified is the scene, which is further categorized into three sub-classes: region, area, and segment. In creating the knowledge base, three main features are identified from the class. They include objects, attributes, and events. These are the key components of the semantic information for the knowledge base. They are discussed below:

- **Objects:** Objects are the entities that represent things that have a distinct existence and can be identified. In this study, the identified objects in the class scene have been categorized into regions, segments, and areas. For this research, we explore and focus on the sub-class, area.
- **Attributes:** Attributes describe the characteristics or properties associated with entities or events. They provide additional information that helps in defining and distinguishing objects or events from one another. Examples include distance, size, and location.
- **Events:** Events refer to actions, occurrences, or happenings that take place around objects. They describe the interactions or changes that occur between objects. Majorly, events help to understand the relationships between different objects. Examples include rain, fire, and storms.

Extracting semantic information such as objects, attributes, and events from the outcome of the object detection model provides a deeper understanding of the meaning and context of the detected objects. It also enables proper knowledge representation. The subclass, area, is further categorized into residential, vegetation, water, and ways.

3.5.2. Ontology Data Storage Implementation and Inferences Ontology Data Storage Implementation

In this research, we employ SPARQL on GraphDB [50] for implementing data storage for the proposed ontology. The sample code below implements an INSERT command for adding triples to the graph store based on triple patterns specified in the ontology design.

Listing 1. SPARQL Query INSERT.

```
PREFIX : <http://deepontos/>
INSERT DATA {
:LowRiseBuilding :IsAdjacentTo :Forest.
:LowRiseBuilding :IsAdjacentTo :HighWay.
:LowRiseBuilding :HasSwimmingPool :Pool.
:FreeWay :IsAdjacentTo :Shoreline }
```

3.5.3. Establishing Knowledge Taxonomy

The knowledge taxonomy provides a foundational framework for organizing, categorizing, and reasoning the knowledge base for the system. It is significant in knowledge representation, semantic modeling, and information retrieval within the ontological system. In this study, the knowledge taxonomy involves the categorization of objects detected based on their shared characteristics and relationships. This provides a structured framework for representing and organizing the knowledge base in a systematic manner. Each class represents a concept, and the hierarchy shows the relationships between these classes. The higher-level classes are more general and encompass broader categories, while the lower-level classes are more specific and represent narrower subcategories.

The description logics employed in this study are based on the subsumption that with a given class, the inferred subclasses can be computed. Therefore, three (3) direct subclasses, region, segment, and area, can be inferred from the main class, scene. Also, the class, area, has four (4) subclasses: residential area, water area, way area, and vegetation area. The subclasses represent narrower subcategories with specific features and properties such as the land cover of the areas (e.g., way area, indicating an area on the scene used as the route, water area, indicating bodies of water that are both natural and man-made, and vegetation area, presenting vegetation type such as forests, grass, and shrubs). Relationships are also established within these subcategories. Based on these subsumptions, further subclasses can be inferred from the already established subclasses. For example, the residential area has four (4) subclasses, the vegetation area has three (3) subclasses, the way area has four (4) subclasses, and the water area has four (4) subclasses. The description logic semantics for the classes and the subsequent subclasses are specified below:

- Scene concept: The Scene concept is represented by the subsumptions in Equations (1)–(3); region, area and Ssgment are represented as subsets of class scene.

$$Region \sqsubseteq Scene \quad (1)$$

$$Area \sqsubseteq Scene \quad (2)$$

$$Segment \sqsubseteq Scene \quad (3)$$

- Area concept: The area class is represented in Equations (4)–(7), as given below:

$$VegetationArea \sqsubseteq Area \sqsubseteq Scene \quad (4)$$

$$WaterArea \sqsubseteq Area \sqsubseteq Scene \quad (5)$$

$$WayArea \sqsubseteq Area \sqsubseteq Scene \quad (6)$$

$$ResidentialArea \sqsubseteq Area \sqsubseteq Scene \quad (7)$$

- Water area concept: The water area class is formed based on the presence of water bodies, both manmade and natural in the area. The water bodies captured in the remotely sensed satellite imagery include rivers, shorelines, pools, lakes, and canals. This is represented in Equations (8)–(12), as given below:

$$\text{River} \sqsubseteq \text{WaterArea} \quad (8)$$

$$\text{Shoreline} \sqsubseteq \text{WaterArea} \quad (9)$$

$$\text{Pool} \sqsubseteq \text{WaterArea} \quad (10)$$

$$\text{Lake} \sqsubseteq \text{WaterArea} \quad (11)$$

$$\text{Canal} \sqsubseteq \text{WaterArea} \quad (12)$$

- Vegetation area concept: The vegetation types captured in the remotely sensed satellite imagery include forests, shrubs, and grasslands. The vegetation area class is represented in Equations (13)–(15):

$$\text{Forest} \sqsubseteq \text{VegetationArea} \quad (13)$$

$$\text{Shrub} \sqsubseteq \text{VegetationArea} \quad (14)$$

$$\text{GrassLand} \sqsubseteq \text{VegetationArea} \quad (15)$$

- Ontology description of the way area concept: As represented in Equations (16)–(19), the way area class, which is a direct subclass of the area class, has the subclasses highway, rural road, local street, and freeway. Likewise, a subclass of the water area class, shoreline, exhibits the property *isAdjacentTo*, in which the freeway subclass is adjacent to the shoreline class, representing closeness to a water body, as given below:

$$\text{Highway} \sqsubseteq \text{WayArea} \quad (16)$$

$$\text{RuralRoad} \sqsubseteq \text{WayArea} \quad (17)$$

$$\text{LocalStreet} \sqsubseteq \text{WayArea} \quad (18)$$

$$\begin{aligned} \text{FreeWay} \sqsubseteq \text{WayArea} \quad \sqcap \\ \exists \text{ isAdjacentTo. Shoreline} \end{aligned} \quad (19)$$

- Ontology description of the residential area concept: As represented in Equations (20)–(23), the residential area class, which is a direct subclass of the area class, has the subclasses of town houses, informal settlements, high-rise buildings, and low-rise buildings. The low-rise buildings subclass has been used to represent the residential areas due to its widespread presence. The shorelines subclass exhibits some features that include adjacency to highways and forests via the property *isAdjacentTo* and some residence instances containing swimming pools. This is presented below:

$$\text{TownHousesBuilding} \sqsubseteq \text{ResidentialArea} \quad (20)$$

$$\text{InformalSettlementsBuilding} \sqsubseteq \text{ResidentialArea} \quad (21)$$

$$\text{HighRiseBuilding} \sqsubseteq \text{ResidentialArea} \quad (22)$$

$$\begin{aligned} \text{LowRiseBuilding} \sqsubseteq \text{ResidentialArea} \quad \sqcap \\ \exists \text{ isAdjacentTo.Highway} \quad \sqcap \\ \exists \text{ isAdjacentTo.Forest} \quad \sqcap \end{aligned} \quad (23)$$

$$\exists \text{ hasSwimmingPools.Pool}$$

4. Results and Discussion

In this section, an analysis of the results of the proposed model is carried out. The performance of the deep learning model is first evaluated using appropriate metrics, such as accuracy, precision, and recall. Sample results of the object detection deep learning model are presented in Figure 5. The results are then analyzed by incorporating feedback into the ontology model for semantic analysis and insights. The ontological reasoning module is composed of a reasoning function that provides further interpretation and analysis of the classification results of the deep learning module based on the ontological reasoning rules. The reasoning function aims to generate inferred information from the taxonomy set and to improve the understanding and enhance the interpretability of the object detection deep learning model output.

4.1. Evaluation Metrics

The deep learning-based object detection techniques were assessed using the proposed dataset. Five metrics, detection accuracy (DA), recall (R), precision (P), average precision (AP), and mean average precision (mAP), were used to analyze the models. Here is a definition of these metrics:

- Detection accuracy: The degree to which a model can accurately identify objects in an image is measured by its detection accuracy. Precision, recall, and F1-score are also measures used to assess a model's detection performance.
- Precision is the percentage of all objects detected by the model that are true positives. With a high precision, the model is less likely to produce false positives, which means that most of the objects it detects are real.
- The average precision of the objects detected by the model is defined as the average precision (AP).
- The overall mean value of the AP is defined as the mean average precision (mAP).
- Recall estimates the percentage of actual objects in the image that the model correctly identified out of all the false positives. A high recall means that the model can successfully identify the majority of the image's objects.

Ontology metrics:

We have also used some metrics to define both the complexity measures and the response time of the proposed ontology model. These metrics are used to quantify the overall design complexity of the ontology and directly infer the reasoning performance using the model's internal structure, among other things. Three primary factors [51] have been considered in the selection of the measures. These include the following:

- Class count: This is the total number of classes in the ontology. It typically provides information on the intricacy and scope of the study's domain covering.
- Class depth: This indicates the degree and depth of hierarchy of the classes and subclasses used in the research.
- Class breadth: This indicates how many subclasses there are on average within a class. It is sometimes referred to as the ontology's branching factor.

The more resources needed to analyze, comprehend, and maintain the ontology, the more complicated the model is, as shown by the greater metric value. As a result, four measures [52] have been chosen to assess the suggested ontology's design complexity. These include the following:

- Number of children (NOC): In the ontological inheritance hierarchy, the NOC represents the total number of its direct offspring. The number of subclasses that are directly derived from a particular class is known as its NOC.
- Depth of inheritance (DIT): In an ontological inheritance hierarchy, DIT is the length of the longest path from a particular class to the root class.
- Class in-degree (CID): In an ontology graph, CID counts the number of edges that point to a node for a given class.

- Class out-degree (CID): Regarding the COD, COD measures the number of edges leaving a given class in the ontology graph.

Reasoning performance: The reasoning performance [53] of the ontology geographical database used in this study in terms of reasoning time and reaction time has been taken into consideration in order to further assess the effectiveness of the suggested ontology:

- Reasoning time: This is the amount of time needed for the reasoning engines used in this study to validate the ontology or infer knowledge.
- Query response time: This gauges how long it typically takes to respond to a given query.

4.2. Class-Wise Detection Performance

The object detection model identified five objects: residences, roads, shorelines, swimming pools, and vegetation from the proposed remotely sensed satellite images dataset. The class-wise detection performance of each of these objects in Table 2 shows that swimming pools were detected with the highest precision score of 62.7%, followed by the vegetation, with a precision score of 57.3%, and shorelines, with a precision score of 54.6%.

Both the residence and roads had comparable detection rates, with residence scoring 41.1 percent, 42.1 percent, 19.3 percent, and 12.8 percent for precision, recall, mAP50, and mAP (50–95), respectively. The roads scored 41.2 percent, 57.1 percent, 13.7 percent, and 4.75 percent for the same metrics. This similarity in detection rates of the objects shows that the objects are evenly distributed within the satellite images captured in the same areas. There are, however, instances of higher detection rates for shorelines, swimming pools, and vegetation in areas where they are densely distributed, as presented in Figure 5a–c. More detection results are presented in Appendix A.



Figure 5. Detecting objects in sample satellite image; (a) Detection of swimming pools and residences; (b) detection of vegetation; (c) detection of residences, roads, and shorelines (Adegun et al. [37]).

Table 2. Class-wise Performance of YOLOv8 on the proposed dataset (Adegun et al. [48]).

Class	P (%)	R (%)	mAP50 (%)	MAP50-95 (%)
Residences	41.1	42.1	19.3	12.8
Roads	41.2	57.1	13.7	4.75
Shorelines	54.6	96.4	99.5	59.7
Swimming pools	62.7	62.9	45.5	12.8
Vegetation	57.3	62.3	12.8	8.45

In Table 3, the performance of the model is compared with some object detection models using precision, recall, mAP50, MAP50-95, and speed.

Table 3. Model’s performance on the proposed dataset (Adegun et al. [48]).

Methods	P (%)	R (%)	mAP50 (%)	MAP50-95 (%)	Speed (ms)
Detectron2 [54]	50	32.7	16	24	0.9
YOLOv5 [55]	53.4	49.7	27	18.4	0.5
YOLOv6 [56]	53.2	47.4	32.1	16.6	0.4
YOLOv7 [57]	54.5	46.2	34.1	25	0.3
YOLOv8 [47]	68	60	43	17.5	0.2

Based on the results in Tables 3–6, the proposed YOLOv8 model has been proven to outperform other object detector methods. This has been established through the following:

- YOLO v8 achieved a very high speed in detecting objects, with the lowest latency rate, as presented in Table 3.
- The performance of YOLOv8 in terms of precision and recall, especially when considering speed, is better than the other objects detectors and more reliable, as presented in Table 3.
- In Table 4, it performs better when compared with some state-of-the-art methods in mAP when experimented on using the publicly available datasets Visdrone and PascalVOC.
- The F1 score and IOU in Table 5 show that the model achieves more than 50% accuracy in detecting most of the objects.
- The confusion matrix in Table 6 also shows that the model is able to achieve a high detection rate for most of the objects correctly.

Table 4. Comparing the model’s performance with state-of-the-art methods.

Authors	Methods	Database	mAP (%)
Xiao et al. [58]	CEM + FPM	PascalVOC	37.2
Koyun et al. [59]	Two-stage object detection framework	VisDrone	42.06
Xu et al. [60]	DetectoRS	AI-TOD	41.9
Zhang et al. [61]	YOLOv5-based SPD	VisDrone-DET2019	41.8
Proposed	YOLOv8-based model	Visdrone	43.1

Table 5. Model’s performance on the proposed dataset based on precision, recall, F1 score, and IOU.

Class	Precision (%)	Recall (%)	F1Score (%)	IOU (%)
Residences	41.1	42.1	41.6	50.6
Roads	41.2	57.1	47.9	58.1
Shorelines	54.6	96	69.5	63.7
Swimming pools	62.7	62.9	62.9	50.1
Vegetation	57.3	62.3	59.7	52.1

Table 6. Model’s performance on the proposed dataset using a confusion matrix.

	Residences	Roads	Shorelines	Swimming Pools	Vegetation
Residence	42	17	10	10	10
Roads	13	57	10	10	10
Shorelines	4	4	96	4	4
Swimming pools	18	15	10	63	10
Vegetation	18	10	10	10	62

4.3. Ontology Model Analysis

In analyzing the proposed ontology model, we used the earlier defined ontology metrics. The NOC value for the scene class is 3 and for the area class is 4 for the ontology that is explained in Section 3. As the NOC value increases, so does the complexity. A higher NOC number also suggests that more subclasses might be impacted by changes made to this class, necessitating more resources in the maintenance of the subclasses. Furthermore, the DIT value for the scene class is 3 and the DIT for the area class is 2, while the DITs for the vegetation area, water area, residential area and way area classes are 1 each for the ontology outlined in Section 3. A higher DIT value indicates that the class reuses more information from its predecessors and is located further down the inheritance tree. A higher DIT value

also suggests that the class is more likely to be impacted by modifications in any of its ancestors, making it more challenging to maintain. The class hierarchy is in a top-down, depth-first fashion, beginning at the root node in the DIT calculation.

Additionally, the CID values for the scene class and area class in the proposed ontology are 3 and 4, respectively. The value of class-in-degree indicates how other nodes are using a certain class. The greater the number of nodes that depend on a certain CID, the higher the value of CID will be. The CID value for the water area class is 5, vegetation area is 3, way area is 4, and residential area is 4. Lastly, the proposed ontology has a COD value of 1 for the vegetation area class and a COD value of 3 for the residential area. The number of nodes that a certain class refers to is indicated by the value of out-degree. In conclusion, this analysis presents a less complex ontology with the maximum DIT of 3 and NOC of 4 in the whole ontology proposed, making it scalable. This will directly affect the reasoning performance in terms of reasoning time and query response time. The output indicates quick reasoning and response time with low computational resources for processing.

4.4. Semantic Reasoning and Event Inferences

The semantic reasoning modules aim to generate inferred information from the taxonomy set based on the ontological reasoning rules. The reasoning rules are established following these logical semantics:

- The symmetric rule stating that if a source instance is related to a target instance, then the target must also be related to the source.
- The transitive rule states that if instance A is related to instance B, and instance B is related to instance C, then A is also related to C.

For instance, let RA be a selected area. The RA area contains all the regions located within the description represented in Equation (24) at the given point. For example, according to the given query, these selected regions are assumed to be residential areas caught in between water areas and way areas during heavy raining events. R_i is a subclass of residential area in way area, and G_i represents a subclass of residential area in the water area, while S represents the spatial relation (a subclass of the property *isAdjacentTo*) of the part of R_i represented by r_i and the part of G_i represented by g_i caught during the event. The query therefore returns the areas that can be influenced by the closeness of the detected objects in the scene captured.

$$RA = \{r_i \in R_i \mid R_i \sqsubseteq ResidentialArea \wedge \exists g_i \in G_i \sqsubseteq ResidentialArea \wedge (r_i, g_i) \in S \sqsubseteq isAdjacentTo\} \quad (24)$$

4.4.1. Ontology Reasoning and Inferences

Three inferential instances drawn from the area class are represented diagrammatically in Figure 6a–c. In Figure 6a, the area class has two subclasses, water area and residential area, whose subclasses swimming pools and low-rise buildings are related through the property *hasSwimmingPools*. The relations between the concepts are provided via two properties; *hasSubclass* and *hasSwimmingPools*. Also in Figure 6b, the Area class has two subclasses VegetationArea and ResidentialArea whose subclasses Forest and LowRising-Building are related through the property *isAdjacentTo*. The relations between the concepts are provided via two properties; *hasSubclass* and *isAdjacentTo*. For the Figure 6c, the Area class has two subclasses WayArea and WaterArea whose subclasses Freeway and Shoreline are related through the property *isAdjacentTo*. The relations between the concepts are provided via two properties; *hasSubclass* and *isAdjacentTo*.

Using SPARQL queries for drawing inferences around the objects in the neighborhood, the outcomes are presented in Figure 7a–c. For example, it can be inferred from Figure 7a that there are some low-rise buildings containing pools, which may be areas of concern. Also, Figure 7b suggests that some low-rise buildings are adjacent to the forest vegetation

type. The outcome from the SPARQL query presents shorelines and freeways, as shown in Figure 7c.

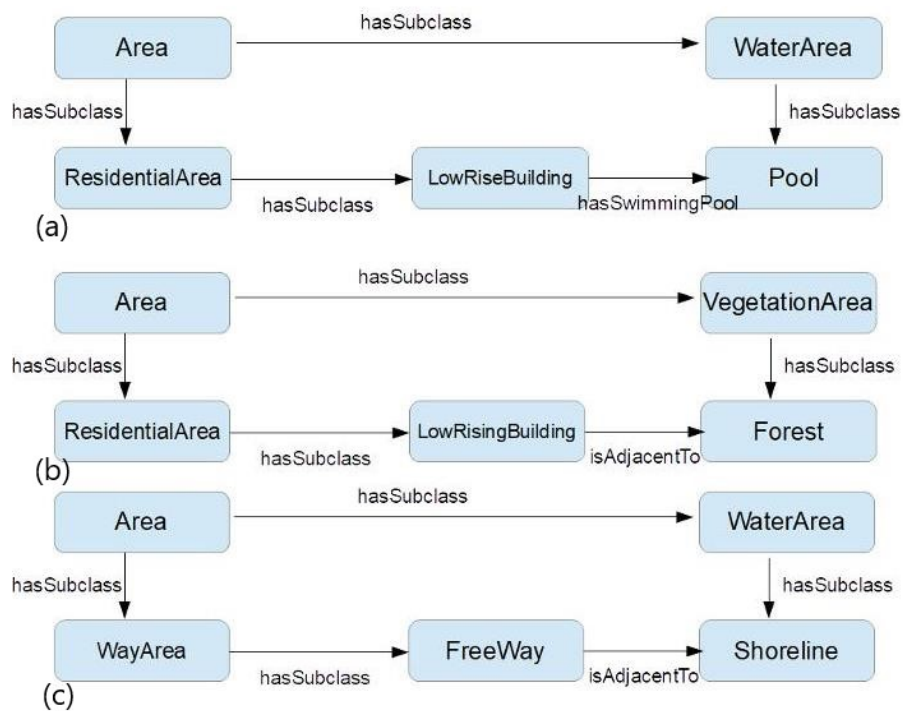


Figure 6. Ontology reasoning with the area class and the subclasses.

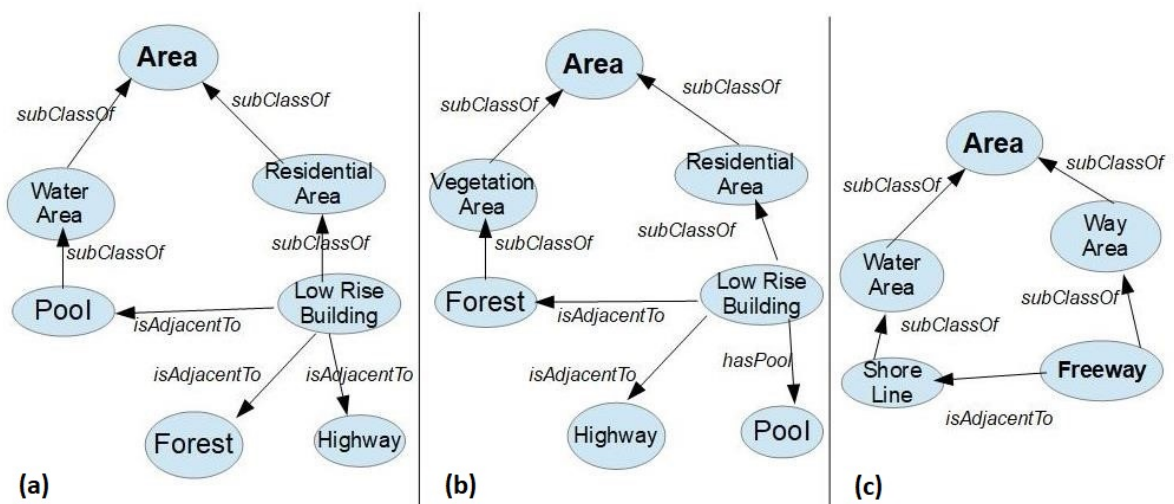


Figure 7. SPARQL outcomes showing ontology reasoning within the area and subclasses.

4.4.2. Visualization Analysis of Detection Results Output

This section presents the visual analysis of the object detection outcomes using the knowledge representation model. In Figure 8a, the objects detected include roads, residences, and vegetation. The detection accuracy of the vegetation is 59%, residences have a detection accuracy of 41%, and roads have a detection accuracy of 51%. This is visualized in Figure 8b, with relations identified including residences adjacent to vegetation, vegetation with a subclass of forest, and way area with a subclass of roads.

In Figure 9a, the objects detected include residence, roads, swimming pools, and vegetation. Swimming pools had the highest detection accuracy of 68%, with residences having 42% and vegetation having 44%. The knowledge representation of the scene is

presented in Figure 9b. The relations established include *isAdjacentTo* and *hasSwimmingPools*. It can be inferred that most of the residences captured had swimming pools with some degree of adjacency to the forest, as shown in Figure 10a. The advanced visualization in Figure 10b shows the presence of a vegetation area, residential area, water area, and way area as the outcomes of the SPARQL query. This also establishes that the residential area was adjacent to the vegetation area and way area with the presence of swimming pools. The Figure 8 output is also further expatiated in Figure 10a using the SPARQL query showing the presence of the vegetation area, residential area, and way area. Relations such as *isAdacentTo* between the subclasses of forests and residences as well as residences and roads are also presented. This outcome presents a more secured area with the presence of the forest.

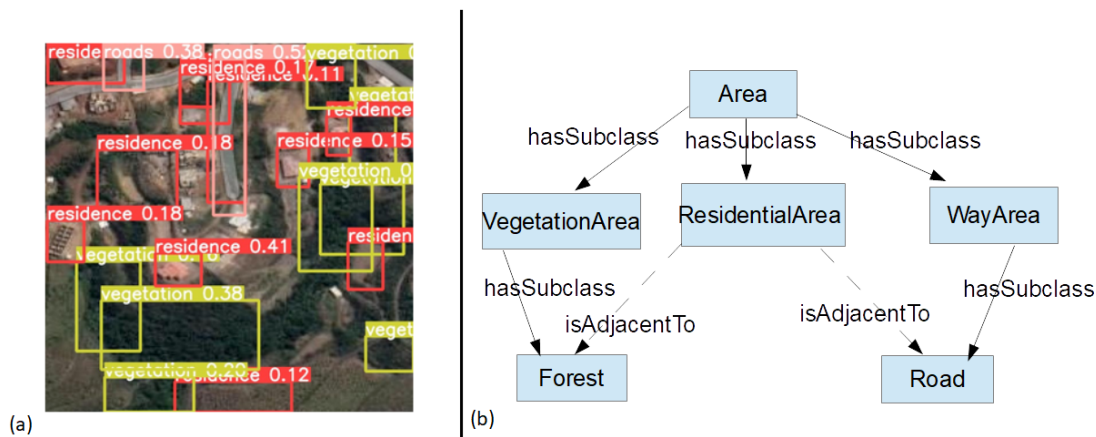


Figure 8. Ontology description of scene classification and object detection output: (a) image showing objects detected, including residences, vegetation, and roads. (b) Local ontology description of detected objects in (a).

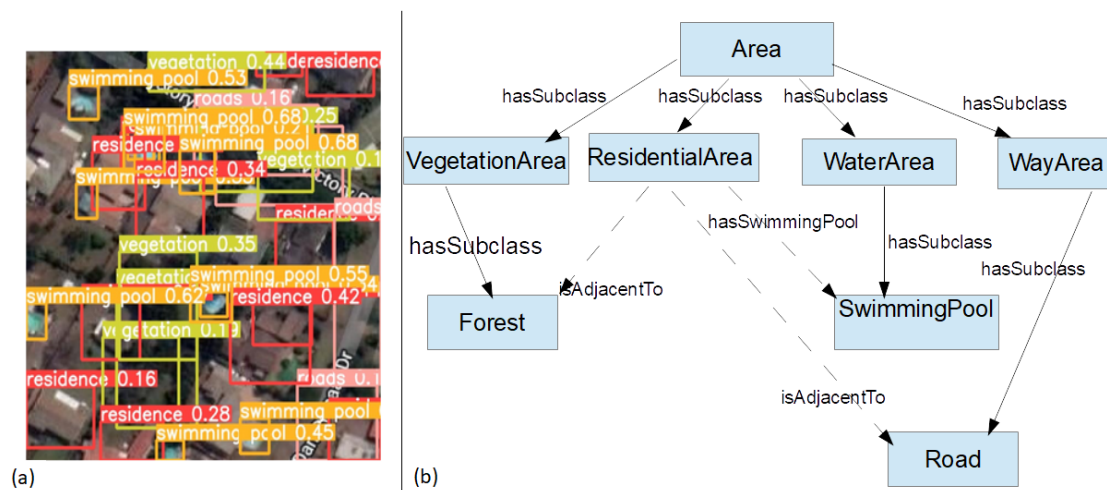


Figure 9. Ontology description of scene classification and object detection output: (a) image showing objects detected including residences, swimming pools, vegetation, and roads. (b) Local ontology description of detected objects in (a).

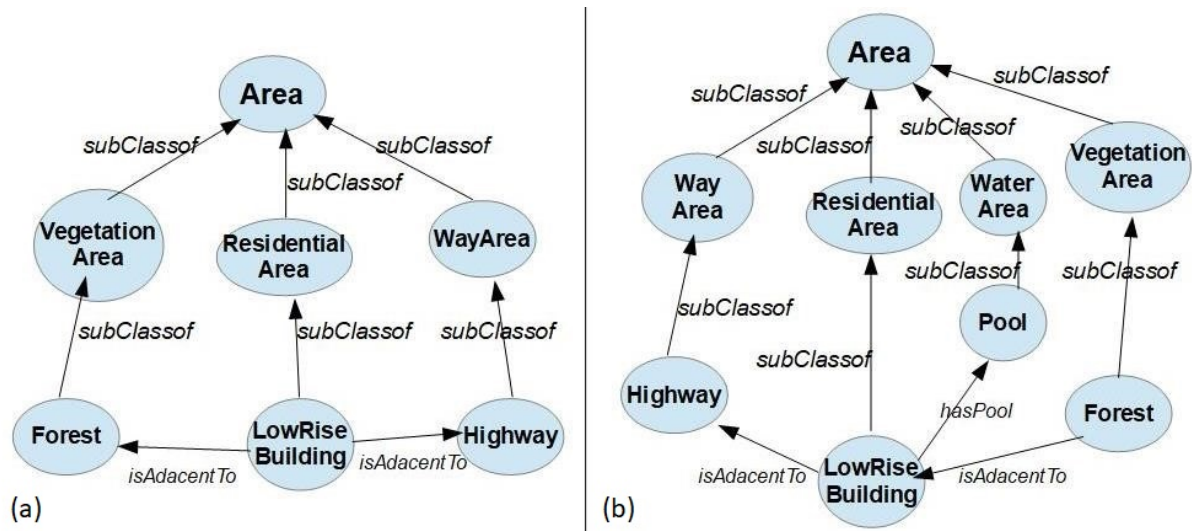


Figure 10. Elaborate description from SPARQL outcome of scene classification and object detection in Figures 8 and 9.

In Figure 11a, the objects detected include roads, swimming pools, residences, vegetation, and shorelines. The detection accuracy of vegetation is 59%, while shorelines have high detection accuracy of 60%, residences have a detection accuracy of 51%, and swimming pools have a detection accuracy of 59%. The ontology description presented in Figure 11b shows the relations identified, including residences with swimming pools, roads adjacent to shorelines, and residences adjacent to vegetation. The presence of shorelines and some residences is conspicuous. Based on the SPARQL query, detailed ontology description of scene classification is produced in Figure 12, showing the presence of vegetation areas, residential areas, water areas, and way areas. It also reveals the presence of shorelines and swimming pools in the water area. The outcome in Figure 12 validates the presence of water areas adjacent to way areas and to residential areas. The result presents a more sensitive location.

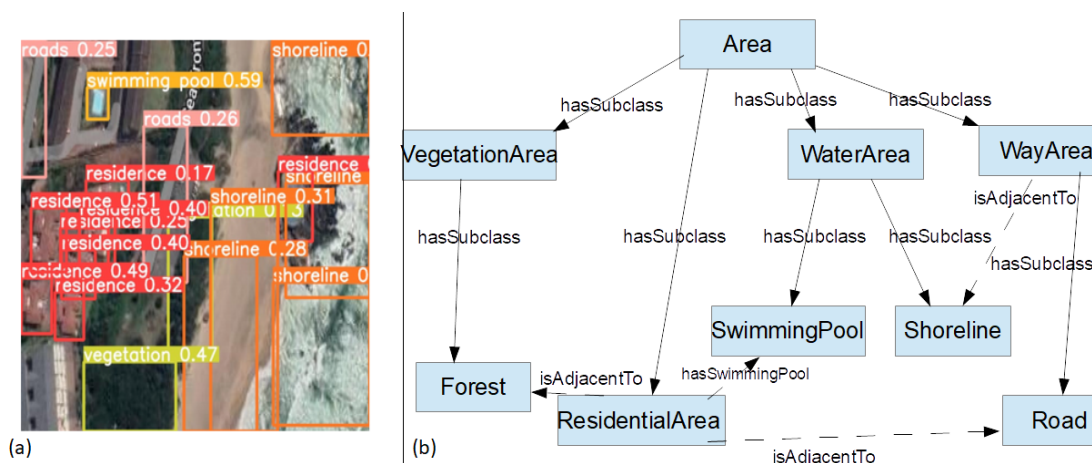


Figure 11. Ontology description of scene classification and object detection output: (a) image showing objects detected including residences, swimming pools, shorelines, vegetation, and roads. (b) Local ontology description of detected objects in (a).

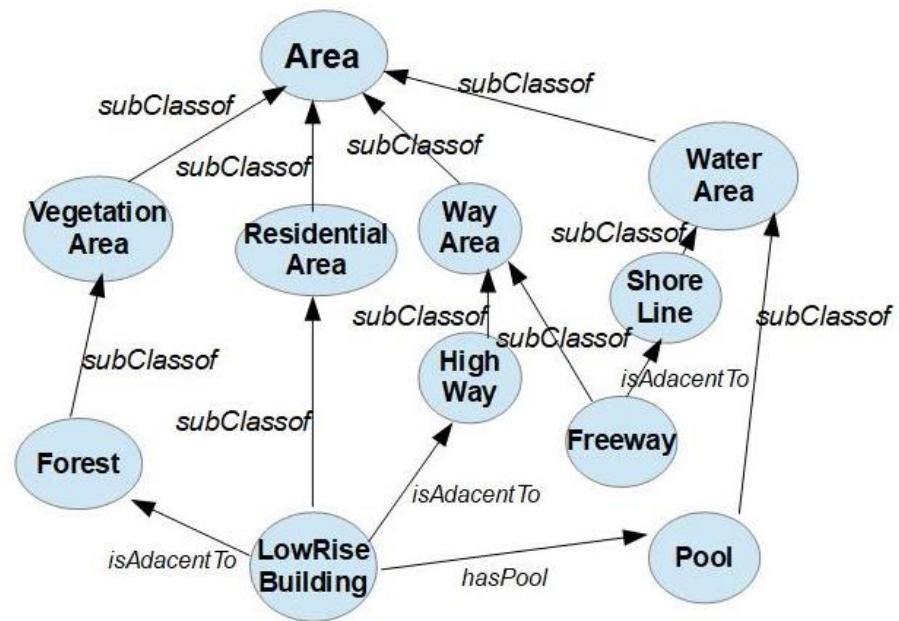


Figure 12. Ontology description based on SPARQL outcome of scene classification and object detection in Figure 11.

4.5. Ontology Modeling—A Case Study of Flood Prevention

Adapting this model specifically to flood prevention, for instance, will lead to providing further queries that can be implemented in the case study. Various instances are highlighted below:

- If a residence is located adjacent to a road, and both the residence and road are affected by flooding, then the residence is at a high risk of flooding. This rule states that if a residence is adjacent to a local road and both the residence and the road are affected by flooding, then the residence is considered to be at a high risk of flooding. The variables “?residence” and “?road” represent the residence and road individuals, respectively. The predicates “Adjacent”, “AffectedByFlooding”, and “HighRiskOfFlooding” represent the relationships between the entities. This rule can be expressed in SWRL [62] as follows:

$$\begin{aligned} & \text{Adjacent}(\text{?residence}, \text{?road}) \wedge \\ & \text{AffectedByFlooding}(\text{?residence}) \wedge \\ & \text{AffectedByFlooding}(\text{?road}) \rightarrow \text{HighRiskOfFlooding}(\text{?residence}) \end{aligned}$$

- If a swimming pool is in a residence, and its drainage system impacts the surrounding area, then it is at high risk of flooding. This rule has the following three conditions: 1. `LocatedIn(?swimmingPool, ?residence)`: The swimming pool is located in a residence. Here, ?swimmingPool represents the individual swimming pool, and ?residence represents the individual residence. 2. `ImpactsSurroundingArea(?drainageSystem drainageSystem)`: The drainage system of the swimming pool has an impact on the surrounding area. ?drainageSystem represents the individual drainage system. The SWRL formulation of this rule is given below.

$$\begin{aligned} & \text{LocatedIn}(\text{?swimmingPool}, \text{?residence}) \wedge \\ & \text{ImpactsSurroundingArea}(\text{?drainageSystem}) \wedge \\ & \text{hasDrainage}(\text{?swimmPool}, \text{?poorDrainage}) \rightarrow \text{HighRiskOfFlooding}(\text{?residence}) \end{aligned}$$

- If vegetation is located on a shoreline, it can help prevent erosion and absorb water. This rule declares that if vegetation is located on a shoreline then it can prevent erosion and absorb water. The variables “?vegetation” and “?shoreline” represent the individual vegetation and shorelines, respectively. The predicates “LocatedOn”, “Can-

PreventErosion”, and “CanAbsorbWater” represent the relationships and properties between the entities. This rule can be expressed in SWRL as follows:

$$\text{LocatedOn}(?vegetation, ?shoreline) \rightarrow \\ \text{CanAbsorbWater}(?vegetation) \wedge \text{CanPreventErosion}(?vegetation)$$

- If a residence is in a low-lying area, then it is at a high risk of flooding. This rule states that if a residence is located in a low-lying area, which is identified by the predicate “LowLyingArea”, then the residence is considered to be at a high risk of flooding. The variables “?residence” and “?lowLyingArea” represent the individual residences and low-lying area, respectively. The predicate “HighRiskOfFlooding” represents the relationship indicating that the residence is at a high risk of flooding. The SWRL representation of this rule is given below.

$$\text{LocatedIn}(?residence, ?lowLyingArea) \wedge \\ \text{LowLyingArea}(?lowLyingArea) \rightarrow \text{HighRiskOfFlooding}(?residence)$$

- If a swimming pool is located uphill from a residence, then it can increase the risk of flooding for the residence by contributing to runoff. This rule states that if a swimming pool is located uphill from a residence and it contributes to runoff, then the residence is considered to be at an increased risk of flooding. The variables “?swimmingPool” and “?residence” represent the swimming pool and residence individuals, respectively. The predicate “LocatedUphill” represents the relationship indicating that the swimming pool is located uphill from the residence. The predicate “ContributesToRunoff” represents the relationship indicating that the swimming pool contributes to runoff. The predicate “IncreasedRiskOfFlooding” represents the relationship indicating that the residence is at an increased risk of flooding. This rule can be represented in SWRL as follows:

$$\text{LocatedUphill}(?swimmingPool, ?residence) \wedge \\ \text{ContributesToRunoff}(?swimmingPool) \rightarrow \text{IncreasedRiskOfFlooding}(?residence)$$

- If a road is adjacent to the residence and located in a floodplain, then it is at high risk of flooding. This rule states that if a road is adjacent to a residence and located in a floodplain, then the road is considered to be at a high risk of flooding. The variables “?road” and “?residence” represent the road and residence individuals, respectively. The predicate “Adjacent” represents the relationship indicating that the road is adjacent to the residence. The predicate “LocatedInFloodplain” represents the relationship indicating that the road is located in a floodplain. The predicate “HighRiskOfFlooding” represents the relationship indicating that the road is at a high risk of flooding. The SWRL representation of this rule is given below.

$$\text{Adjacent}(?road, ?residence) \wedge \\ \text{LocatedInFloodplain}(?road) \rightarrow \text{HighRiskOfFlooding}(?road)$$

- If vegetation is adjacent to the residence, then it can help to absorb rainfall and prevent runoff. This rule states that if vegetation is adjacent to a residence, then it is considered to help absorb rainfall and prevent runoff. The variables “?vegetation” and “?residence” represent the vegetation and residence individuals, respectively. The predicates “Adjacent”, “HelpsAbsorbRainfall”, and “HelpsPreventRunoff” represent the relationships and properties between the entities. This rule can be expressed in SWRL as follows:

$$\text{Adjacent}(?vegetation, ?residence) \rightarrow \\ \text{HelpsAbsorbRainfall}(?vegetation) \wedge \text{HelpsPreventRunoff}(?vegetation)$$

- If a shoreline lacks vegetation, then it is at a high risk of flooding due to storm surges or high tides. This rule states that if a shoreline lacks vegetation, it is considered to be at a high risk of flooding from storm surges or high tides. The variable “?shoreline” represents the individual shoreline. The predicate “LacksVegetation” represents the property indicating that the shoreline lacks vegetation. The predicate “High-

RiskOfFlooding” indicates that the shoreline is at a high risk of flooding. The SWRL representation of this rule is given below.

$$\text{LacksVegetation}(?shoreline) \wedge \text{HighRiskOfFlooding}(?shoreline)$$

4.6. Ontology Model Justification

The incorporation of ontology into deep learning models for object recognition and image classification in the context of smart cities presents several benefits with regard to overall system performance, scalability, interpretability of the models, semantic comprehension, and data integration. These advantages support the use of ontology-based techniques in creating reliable and successful smart city applications in the following areas:

- **Contextual awareness:** The proposed model is able to comprehend the context and semantics of objects inside the smart city environment because the ontologies offer an organized framework for defining and relating concepts.
- **Expandable and scalable knowledge structure:** As smart cities develop, the proposed ontology provides a scalable structure to embrace new ideas and connections.
- **Adaptability to new situations:** The model’s easy extension and adaption to new situations and applications within the context of smart cities is made possible by the structured knowledge representation found in the ontologies.

5. Summary, Future Works, and Conclusions

5.1. Summary

Two major approaches are adopted in addressing the challenges identified in this study. They are as follows:

- **Efficient object detection and scene classification in satellite images.** This is achieved through a robust deep learning architecture utilizing the YOLOv8 model.
- **Ontological modeling for knowledge extraction and inferences.** The ontological model serves as a framework for knowledge extraction and inferences. It enables the capture, organization, and representation of knowledge in a structured and semantically rich manner. With the ontological model in place, it becomes possible to extract relevant knowledge from the knowledge base and make logical inferences based on the relationships and dependencies defined within the model.

5.2. Future Works

This research can integrate deep learning object detection algorithms with ontology to detect objects from remotely sensed satellite images and establish the relationships that exist among the objects and identify potential disaster events. This research solely considers the attributes of the objects to determine identifiable logical rules for developing the ontology model. However, this research does not consider some features such as location, the distance between entities, and the size of the objects to determine the feasibility of disaster. This will be considered in our future research. Also, the research relies on the detection accuracy of the YOLOv8-based objects detection model. The detection accuracy will also be improved in future research, especially in images with the presence of occlusions.

Specifically, the following suggestions are considered as the future directions of this research:

- **Actual historical data of some environments will be used in the future works to demonstrate how the ontological model will predict flooding in the environment.** The ontological model will also be evaluated against the ground truths.
- **Develop techniques to utilize the semantic context provided by the ontology reasoning module to refine object detection results.**
- **Explore methods to fuse information from different modalities, such as textual descriptions or sensor data, with object detection results and use ontology reasoning to correlate and enhance the understanding of the combined data.**

- Extend the integration to handle temporal aspects by incorporating ontology reasoning to track and reason about object interactions and changes over time.
- Explore ways to customize the integration for specific application domains, such as autonomous vehicles, healthcare, or industrial automation, and tailor the ontology reasoning and object detection integration to the unique requirements of each domain.

5.3. Conclusions

This study proposes an ontology-based deep learning framework for the analysis and classification of remotely sensed satellite images. The framework integrates object detection deep learning models with an ontology model to provide semantic analysis and interpretability of the objects detected in these images. The experiments conducted on newly created datasets, captured from high-resolution remote sensing satellite images, demonstrate the effectiveness of the proposed model. The system employs a YOLOv8-based deep learning model for object detection, and it achieved promising results of a 68% precision score and 60% recall score when evaluated using the proposed dataset. The system also employs a less complex ontology model with a maximum DIT of 3 and NOC of 4 in the whole network. This study highlights the potential of this framework for accurate object detection, semantic analysis, and its applicability in various domains, including environmental monitoring and disaster management such as floods. Further research and refinements will be made in future works to enhance the performance and broaden the scope of this model for real-world applications.

Author Contributions: Conceptualization, A.A.A. and J.V.F.-D.; methodology, A.A.A.; software, A.A.A.; validation, A.A.A., J.V.F.-D., S.V., and J.O.; formal analysis, A.A.A.; investigation, A.A.A., J.V.F.-D.; resources, J.V.F.-D., S.V. and J.O.; data curation, A.A.A.; writing—original draft preparation, A.A.A., J.V.F.-D., and S.V.; writing—review and editing, A.A.A. and J.V.F.-D.; visualization, A.A.A.; supervision, J.V.F.-D., S.V., and J.O.; project administration, J.V.F.-D.; funding acquisition, J.V.F.-D. and S.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The datasets generated during and/or analyzed during the current study are publicly available at: <https://data.mendeley.com/datasets/s5v4zz7yj5/1> and are also presented in the form of tables and figures in this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

More results showing the detection of smaller objects, as seen from RS images such as roads, townhouses, etc.



References

1. Ennouri, K.; Smaoui, S.; Triki, M.A. Detection of Urban and Environmental Changes via Remote Sensing. *Circ. Econ. Sustain.* **2021**, *1*, 1423–1437. [[CrossRef](#)] [[PubMed](#)]
2. Basheer, S.; Wang, X.; Farooque, A.A.; Nawaz, R.A.; Liu, K.; Adekanmbi, T.; Liu, S. Comparison of land use land cover classifiers using different satellite imagery and machine learning techniques. *Remote Sens.* **2022**, *14*, 4978. [[CrossRef](#)]
3. Gao, H.; Guo, J.; Guo, P.; Chen, X. Classification of very-high-spatial-resolution aerial images based on multiscale features with limited semantic information. *Remote Sens.* **2021**, *13*, 364. [[CrossRef](#)]
4. Zhao, W.; Bo, Y.; Chen, J.; Tiede, D.; Blaschke, T.; Emery, W.J. Exploring semantic elements for urban scene recognition: Deep integration of high-resolution imagery and OpenStreetMap (OSM). *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 237–250 [[CrossRef](#)]
5. Ma, A.; Wan, Y.; Zhong, Y.; Wang, J.; Zhang, L. SceneNet: Remote sensing scene classification deep learning network using multi-objective neural evolution architecture search. *ISPRS J. Photogramm. Remote Sens.* **2021**, *172*, 171–188. [[CrossRef](#)]
6. Kaur, R.; Singh, S. A comprehensive review of object detection with deep learning. *Digit. Signal Process.* **2022**, 103812. [[CrossRef](#)]

7. Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Deep learning for hyperspectral image classification: An overview. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6690–6709. [[CrossRef](#)]
8. Linardos, V.; Drakaki, M.; Tzionas, P.; Karnavas, Y.L. Machine learning in disaster management: Recent developments in methods and applications. *Mach. Learn. Knowl. Extr.* **2022**, *4*, 446–473. [[CrossRef](#)]
9. Husni, N.; Latifah; Sari, P.A.R.; Handayani, A.S.; Dewi, T.; Seno, S.A.H.; Caesarendra, W.; Glowacz, A.; Oprzędkiewicz, K.; Sułowicz, M. Real-time littering activity monitoring based on image classification method. *Smart Cities* **2021**, *4*, 1496–1518. [[CrossRef](#)]
10. Peixoto, J.P.J.; Costa, D.G.; Portugal, P.; Vasques, F. Flood-Resilient Smart Cities: A Data-Driven Risk Assessment Approach Based on Geographical Risks and Emergency Response Infrastructure. *Smart Cities* **2024**, *7*, 662–679. [[CrossRef](#)]
11. Shokri, D.; Larouche, C.; Homayouni, S. A comparative analysis of multi-label deep learning classifiers for real-time vehicle detection to support intelligent transportation systems. *Smart Cities* **2023**, *6*, 2982–3004. [[CrossRef](#)]
12. Nadeem, M.; Dilshad, N.; Alghamdi, N.S.; Dang, L.M.; Song, H.-K.; Nam, J.; Moon, H. Visual Intelligence in Smart Cities: A Lightweight Deep Learning Model for Fire Detection in an IoT Environment. *Smart Cities* **2023**, *6*, 2245–2259. [[CrossRef](#)]
13. Thakker, D.; Mishra, B.K.; Abdullatif, A.; Mazumdar, S.; Simpson, S. Explainable artificial intelligence for developing smart cities solutions. *Smart Cities* **2020**, *3*, 1353–1382. [[CrossRef](#)]
14. Cong, Y.; Inazumi, S. Integration of Smart City Technologies with Advanced Predictive Analytics for Geotechnical Investigations. *Smart Cities* **2024**, *7*, 1089–1108. [[CrossRef](#)]
15. Chen, X.; Jiang, N.; Li, Y.; Cheng, G.; Liang, Z.; Ying, Z.; Zhang, Q.; Zhao, R. Efficient Decoder and Intermediate Domain for Semantic Segmentation in Adverse Conditions. *Smart Cities* **2024**, *7*, 254–276. [[CrossRef](#)]
16. Li, W.; Chen, K.; Chen, H.; Shi, Z. Geographical knowledge-driven representation learning for remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5405516. [[CrossRef](#)]
17. Moran, N.; Nieland, S.; Kleinschmit, B. Combining machine learning and ontological data handling for multi-source classification of nature conservation areas. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *54*, 124–133. [[CrossRef](#)]
18. Li, Y.; Ouyang, S.; Zhang, Y. Combining deep learning and ontology reasoning for remote sensing image semantic segmentation. *Knowl. Based Syst.* **2022**, *243*, 108469. [[CrossRef](#)]
19. Li, W.; Zhou, X.; Wu, S. An integrated software framework to support semantic modeling and reasoning of spatiotemporal change of geographical objects: A use case of land use and land cover change study. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 179. [[CrossRef](#)]
20. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. Joint Deep Learning for land cover and land use classification. *Remote Sens. Environ.* **2019**, *221*, 173–187. [[CrossRef](#)]
21. Potnis, A.V.; Durbha, S.S.; Shinde, R.C. Semantics-driven remote sensing scene understanding framework for grounded spatio-contextual scene descriptions. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 32. [[CrossRef](#)]
22. Xie, X.; Zhou, X.; Li, J.; Dai, W. An ontology-based framework for complex urban object recognition through integrating visual features and interpretable semantics. *Complexity* **2020**, *2020*, 5125891. [[CrossRef](#)]
23. Bouyerbou, H.; Bechkoum, K.; Benblidia, N.; Lepage, R. Ontology-based semantic classification of satellite images: Case of major disasters. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2347–2350
24. Almendros-Jimenez, J.M.; Domene, L.; Piedra-Fernandez, J.A. A framework for ocean satellite image classification based on ontologies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *6*, 1048–1063. [[CrossRef](#)]
25. Alirezaie, M.; Kiselev, A.; Långkvist, M.; Klügl, F.; Loutfi, A. An ontology-based reasoning framework for querying satellite images for disaster monitoring. *Sensors* **2017**, *17*, 2545. [[CrossRef](#)] [[PubMed](#)]
26. Filali, J.; Zghal, H.B.; Martinet, J. Ontology-based image classification and annotation. *Int. J. Pattern Recognit. Artif. Intell.* **2020**, *34*, 2040002. [[CrossRef](#)]
27. Wang, C.; Zhuo, X.; Li, P.; Chen, N.; Wang, W.; Chen, Z. An ontology-based framework for integrating remote sensing imagery, image products, and in situ observations. *J. Sens.* **2020**, *2020*, 1–12. [[CrossRef](#)]
28. Fang; Weili; Ma, L.; Love, P.E.D.; Luo, H.; Ding, L.; Zhou, A.O. Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology. *Autom. Constr.* **2020**, *119*, 103310. [[CrossRef](#)]
29. Miranda, E.; Mutiara, A.B.; Ernastuti, E.; Wibowo, W.C. Land Cover Classification through Ontology Approach from Sentinel-2 Satellite Imagery. *Int. J. Geoinformatics* **2020**, *16*, 61–72.
30. Sambandam, P.; Yuvaraj, D.; Padmakumari, P.; Swaminathan, S. Deep attention based optimized Bi-LSTM for improving geospatial data ontology. *Data Knowl. Eng.* **2023**, *144*, 102123. [[CrossRef](#)]
31. Zhang, Y.; Qu, Y.; Li, C.; Lei, Y.; Fan, J. Ontology-driven hierarchical sparse coding for large-scale image classification. *Neurocomputing* **2019**, *360*, 209–219. [[CrossRef](#)]
32. Benkirane, F.E.; Crombez, N.; Ruichek, Y.; Hilaire, V. Integration of ontology reasoning-based monocular cues in deep learning modeling for single image depth estimation in urban driving scenarios. *Knowl.-Based Syst.* **2023**, *260*, 110184. [[CrossRef](#)]
33. Kwenda, C.; Gwetu, M.; Fonou-Dombeu, J.V. Ontology with Deep Learning for Forest Image Classification. *Appl. Sci.* **2023**, *13*, 5060. [[CrossRef](#)]
34. Gupta, U.; Chaudhury, S. Deep transfer learning with ontology for image classification. In Proceedings of the 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Patna, India, 16–19 December 2015; pp. 1–4.

35. Kim, S.; Jeon, T.H.; Rhiu, I.; Ahn, J.; Im, D.-H. Semantic scene graph generation using RDF model and deep learning. *Appl. Sci.* **2021**, *11*, 826. [CrossRef]
36. Li, W.; Zhang, L.; Wu, C.; Cui, Z.; Niu, C. A new lightweight deep neural network for surface scratch detection. *Int. J. Adv. Manuf. Technol.* **2022**, *123*, 1999–2015. [CrossRef] [PubMed]
37. Chandra, R.; Agarwal, S.; Singh, N. Semantic sensor network ontology based decision support system for forest fire management. *Ecol. Inform.* **2022**, *72*, 101821. [CrossRef]
38. Li, Y.; Wei, H.; Han, Z.; Jiang, N.; Wang, W.; Huang, J. Computer vision-based hazard identification of construction site using visual relationship detection and ontology. *Buildings* **2022**, *12*, 857. [CrossRef]
39. Patel, A.S.; Merlino, G.; Puliafito, A.; Vyas, R.; Vyas, O.P.; Ojha, M.; Tiwari, V. An NLP-guided ontology development and refinement approach to represent and query visual information. *Expert Syst. Appl.* **2023**, *213*, 118998. [CrossRef]
40. Vonderhaar, L.; Elvira, T.; Procko, T.; Ochoa, O. Towards Robust Training Datasets for Machine Learning with Ontologies: A Case Study for Emergency Road Vehicle Detection. *arXiv* **2024**, arXiv:2406.15268.
41. Raj, N.P.; Tarun, G.; Santosh, D.T.; Raghava, M. Ontological Scene Graph Engineering and Reasoning Over YOLO Objects for Creating Panoramic VR Content. In Proceedings of the International Conference on Multi-disciplinary Trends in Artificial Intelligence, Hyderabad, India, 21–22 July 2023; Springer Nature: Cham, Switzerland, 2023; pp. 225–235.
42. Mohan, A.; Singh, A.K.; Kumar, B.; Dwivedi, R. Review on remote sensing methods for landslide detection using machine and deep learning. *Trans. Emerg. Telecommun. Technol.* **2021**, *32*, e3998. [CrossRef]
43. Ouchra, H.; Belangour, A. Satellite image classification methods and techniques: A survey. In Proceedings of the 2021 IEEE International Conference on Imaging Systems and Techniques (IST), Kaohsiung, Taiwan, 24–26 August 2021; pp. 1–6.
44. Kumar, L.; Mutanga, O. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sens.* **2018**, *2*, 1509. [CrossRef]
45. Google Earth Engine. 2023. Available online: <https://earthengine.google.com/> (accessed on 1 June 2024).
46. Roboflow. 2022. Available online: <https://roboflow.com/> (accessed on 1 June 2024).
47. Solawetz, J. Francesco, “What is YOLOv8? The Ultimate Guide”. 2023-1-11 (2023). Available online: <https://blog.roboflow.com/whats-new-in-yolov8/> (accessed on 1 June 2024).
48. Adegun, A.A.; Dombou, J.V.F.; Viriri, S.; Odindi, J. State-of-the-Art Deep Learning Methods for Objects Detection in Remote Sensing Satellite Images. *Sensors* **2023**, *23*, 5849. [CrossRef]
49. Glimm, B. Using SPARQL with RDFS and OWL entailment. In *Reasoning Web International Summer School*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 137–201.
50. Ontotext. 2024. Available online: <https://www.ontotext.com/> (accessed on 1 June 2024).
51. Al-Hassan, M.; Abu-Salih, B.; Al Hwaitat, A. DSpamOnto: An Ontology Modelling for Domain-Specific Social Spammers in Microblogging. *Big Data Cogn. Comput.* **2023**, *7*, 109. [CrossRef]
52. Zhang, H.; Li, Y.-F.; Tan, H.B.K. Measuring design complexity of semantic web ontologies. *J. Syst. Softw.* **2010**, *83*, 803–814. [CrossRef]
53. Casini, G.; Meyer, T.; Moodley, K.; Varzinczak, I. *Towards Practical Defeasible Reasoning for Description Logics*; Centre for Artificial Intelligence Research: Hong Kong, China, 2013.
54. Merz, G.; Liu, Y.; Burke, C.J.; Aleo, P.D.; Liu, X.; Kind, M.C.; Kindratenko, V.; Liu, Y. Detection, instance segmentation, and classification for astronomical surveys with deep learning (DEEPDISC): DETECTRON2 implementation and demonstration with Hyper Suprime-Cam data. *Mon. Not. R. Astron. Soc.* **2023**, *526*, 1122–1137. [CrossRef]
55. Jocher, G.; Stoken, A.; Borovec, J.; Chaurasia, A.; Changyu, L.; Hogan, A.; Hajek, J.; Diaconu, L.; Kwon, Y.; Defretin, Y.; et al. ultralytics/yolov5: v5.0-YOLOv5-P6 1280 models, AWS, Supervise.ly and YouTube integrations. Zenodo 2021. Available online: <https://github.com/ultralytics/yolov5/releases/tag/v5.0> (accessed on 24 July 2024).
56. Li, C.; Li, L.; Jiang, H.; Weng, K.; Geng, Y.; Li, L.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W.; et al. YOLOv6: A single-stage object detection framework for industrial applications. *arXiv* **2022**, arXiv:2209.02976.
57. Wang, C.-Y.; Bochkovskiy, A.; Mark, H.-Y. Liao. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 7464–7475.
58. Xiao, J.; Guo, H.; Zhou, J.; Zhao, T.; Yu, Q.; Chen, Y.; Wang, Z. Tiny object detection with context enhancement and feature purification. *Expert Syst. Appl.* **2023**, *211*, 118665. [CrossRef]
59. Koyun, O.C.; Keser, R.K.; Akkaya, I.B.; Töreyin, B.U. Focus-and-Detect: A small object detection framework for aerial images. *Signal Process Image Commun.* **2022**, *104*, 116675. [CrossRef]
60. Xu, C.; Wang, J.; Yang, W.; Yu, H.; Yu, L.; Xia, G.-S. Detecting tiny objects in aerial images: A normalized Wasserstein distance and a new benchmark. *ISPRS J. Photogramm. Remote Sens.* **2022**, *190*, 79–93. [CrossRef]

61. Zhang, J.; Wan, G.; Jiang, M.; Lu, G.; Tao, X.; Huang, Z. Small object detection in UAV image based on improved YOLOv5. *Syst. Sci. Control. Eng.* **2023**, *11*, 2247082. [[CrossRef](#)]
62. O'Connor, M.; Musen, M.; Das, A. Using the semantic web rule language in the development of ontology-driven applications. In *Handbook of Research on Emerging Rule-Based Languages and Technologies: Open Solutions and Approaches*; IGI Global: Hershey, PA, USA, 2009; pp. 525–539.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.