

Article

Not peer-reviewed version

---

# Distributed Artificial Intelligence for Organizational and Behavioral Recognition of Bees and Ants

---

[Apolinar Velarde Martínez](#)<sup>\*</sup> and [Gilberto González Rodríguez](#)

Posted Date: 17 December 2025

doi: 10.20944/preprints202512.1510.v1

Keywords: scene graph generation; graph neural network; behavioral recognition; Internet of Things; sensors; artrophods



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

# Distributed Artificial Intelligence for Organizational and Behavioral Recognition of Bees and Ants

Apolinar Velarde Martinez \*  and Gilberto Gonzalez Rodriguez

Instituto Tecnológico el Llano Aguascalientes, Mexico

\* Correspondence: apolinar.vm@llano.tecnm.mx

## Abstract

Scientific studies have demonstrated how certain insect species can be used as bioindicators and reverse environmental degradation through their behavior and organization. Studying these species involves capturing and extracting hundreds of insects from a colony for subsequent study, analysis, and observation. This allows researchers to classify the individuals and also determine the organizational structure and behavioral patterns of the insects within colonies. The miniaturization of hardware devices for data and image acquisition, coupled with new Artificial Intelligence techniques such as Scene Graph Generation (SGG), has evolved from the detection and recognition of objects in an image to the understanding of relationships between objects and the ability to produce textual descriptions based on image content and environmental parameters. This research paper presents the design and functionality of a distributed computing architecture for image and video acquisition of bees and ants in their natural environment, and a parallel computing architecture that hosts two datasets with images of real environments from which scene graphs are generated to recognize, classify, and analyze the behavior of bees and ants, while preserving and protecting these species. The experiments carried out are classified into the recognition and classification of objects in the image, as well as the understanding of the relationships between objects and the generation of textual descriptions of the images. The results of the experiments, conducted in real-life environments, show recognition rates above 70%, classification rates above 80%, and comprehension and generation of textual descriptions with an assertive rate of 85%.

**Keywords:** scene graph generation; graph neural network; behavioral recognition; Internet of Things; sensors; arthropods

---

## 0. Introducción

Scientific research, through studies and observations, has demonstrated the existence of beneficial species that can reverse environmental degradation caused by industrial development and human overpopulation. Ants and bees are species that restore the environment [14], and serve as biological indicators of environmental change [2]. Biologists, entomologists, behavioral neurobiologists, and behavioral ecologists [3] have studied and observed behaviors such as conduct and organization in these species. However, although these insects have intrigued specialists for many years due to their peculiar behavior and organizational structure, many questions remain unanswered, and studies on them continue with the goal of understanding the beneficial impacts they provide to the planet.

The benefits provided by these arthropods are many and varied. Ants are ecological and environmental indicators in ecological restoration and rehabilitation processes [9], they constantly remove particles from the substrate where they live, favoring the flow of nutrients and their mineralization [6], also actively participate in seed dispersal [7,8], offer better benefits in biological pest control as opposed to pesticides, which have no effect on some pest species that have developed resistance to pesticides [14] and are used to predict the conditions of environmental conservation [5].

In the same way, honeybees are pollinators in intensive agricultural systems [18–20], they produce foods with high nutritional value [15], such as honey and propolis, and they are also producers of pollen and wax, [16]; *apis mellifera* colonies are used as bioindicators as they allow environmental sampling of different nature [21]; other research works show that honey produced by bees can also be used as a bioindicator [16,17].

Studies targeting these insects involve trapping hundreds of insects from a colony, which are then subjected to containers of poisonous liquids, which cause harm or death when ingested, inhaled, or absorbed, causing their normal practices to be violated. To conduct ant studies in ant colonies, pitfall traps distributed in transects are used; pitfall traps consist of 100 mL plastic containers placed flush with the ground and partially filled with a 30% propylene glycol solution and a few drops of detergent to retain and preserve the intercepted ants; collected specimens are preserved in 96% ethanol [5,8]. In the same way, to conduct bee studies, two types of sampling are carried out: double sieve sampling and powdered sugar sampling [22]; these samples imply opening of hives, collecting hundreds of bees to be sacrificed with isopropyl alcohol and discarded with filters to verify the health status of the hive [23,26,27]. Another study [29] reports that bees are collected on the same day of the experiment, anesthetized, and briefly cooled in ice. The vast majority of studies are carried out in distributed sites, in order to obtain the diversity of species and different sites where the species live.

Each member of an ant or bee colony has specific roles to play. Some are workers, others drones, and the queen of the colony is the only one. The disappearance of individuals can jeopardize the organizational capacity of these arthropods and place the colony on the brink of collapse. For example, queens are insects that represent the entire population of the nest, so their murder can put the entire colony on the brink of collapse. In certain ant species, smaller ants protect the food they find, while larger ones transport it back to the nest. Others are foraging workers, others care for the young. Some specialize in collecting pollen, water, or tree resin, while others specialize in pheromone trails [30].

Technological advances in computer vision have gone from detecting and recognizing objects in an image, to understanding the relationships between objects and detecting a visual relationship between them, to generating a textual description and creating a caption of the image based on the image content [1]. Scene graphs are a structured representation of a scene that can express the objects, attributes and social relationships between objects in the scene. Also, the miniaturization of devices with next-generation hardware operated with solar-powered electric devices can access remote sites to capture, store, process and output data analysis results.

Considering the above described and taking into account the advances in deep learning, and the generation of scene graphs to recognize objects, attributes and relationships between objects in the scene, it is possible to build a non-invasive artificial vision computational system to recognize, classify and describe the organizational and behavioral behavior of insects, while preserving and avoiding their mass sacrifice.

To observe organizational and behavioral characteristics, nest access is a little-explored location in research targeting ants, and in bee research, some work has focused on the beehive entrance [24,25]; some research records bee traffic and flower visits at the beehive entrance, providing interesting information on how several bees interact to take advantage of a common set of flowers. In this research, we consider nest access to be the best observation location because it is where all the inhabitants of that nest must go. They come out for the first time after birth, it is the space through which food is brought closer, and it is the best-guarded location to prevent intruders from entering. Bee traffic at the nest entrance provides interesting information on how several bees interact to take advantage of a common set of flowers [32]. Thus, with a system installed at the nest entrance, it is possible to obtain information on distinctive life history traits and show distinctive responses to natural and anthropogenic disturbances.

To do this, two tools from the computational field are used: distributed computing and parallel computing; *first*, for distributed computing we use miniaturized devices from the Internet of Things (IoT) with which images and videos of insects are obtained from remote locations; *second*, with high-

performance parallel computing we can design Distributed Artificial Intelligence, that allows us to carry out automatic shape recognition and obtain knowledge of the behavior of living beings in uncontrolled environments. In addition to the above, in the extensive literature reviewed, a system that uses the miniaturization of computing devices and Artificial Intelligence techniques to conduct studies such as those proposed in this research, related to bees and ants, was not found. This paper is an extension version of the conference paper [28].

This research work is organized as follows: In the 1 section, a set of works related to the study and observation of ants and bees, as well as applications of Scene Graph Generation, are described. The general objective and specific objectives of the project are listed in the 2 section. The justification of the project is in the 3 section. A set of definitions related to this work appears in the 4 section; the formal definition of the problem is in the 5 section; the materials and methods are in the 6 section. The Architecture of Distributed Artificial Intelligence is in the 7 section. The experiments are described in the 8 section. Finally, conclusions and discussion are presented in the 9 and 10 sections respectively.

## 1. Related Works

In this section, a set of studies found in the literature related to scene graph generation are described. These studies have been used in this research project. The section is organized into three subsections: first, we present the studies related to ant studies in different countries, which highlight the importance of ants as bioindicators and the parameters that drive ant behavior; second, we present the studies related to bee behavior; and finally, the studies related to scene graph generation.

### 1.1. Ants

In [14] aims to understand the net effects of ants an biological control considering their services and disservices, hypothesized that field size, crop system and experiment duration would modulate the effects of ants on the abundance of pests and their natural enemies, plant damage and crop yield; three variables were used, abundance of natural enemies, plant damage and crop yield. The crops studied were different, the most abundant being citrus crops (169 cases), mango (22), apple and cocoa (21 cases each). The characteristic observed in the ants was their size; extracted information about the most abundant ant species in each study. they used the Global Ant Database [46] and additional literature to assess antbody length. They found that the size of the most abundant ant is a proxy for one important trait of this organism among many others.

The study conducted by [5] aimed to evaluate the bioindicator potential of the ant fauna. This study aimed a) to characterize the diversity of ants in different habitat strata in the Parque Estadual do Turvo located in Derrubadas, Rio Grande do Sul, Brazil and b) to analyze the bioindicator potential of ant species for the soil and leaf litter and arboreal strata; sampled the ant assemblage at five sites distributed along the road that crosses the park to the Uruguay River. Sampling methods were soil and canopy pitfall traps, sardine baits, glucose baits, entomological umbrellas, and a sweeping net. Recorded a richness of 157 species belonging to 32 genera and eight subfamilies. Observed that only nine species (5.7% of the sampled richness) had a significant Indication Value. The authors of the research consider that ant species can bioindicate soil cover and the presence of leaf litter, vegetation diversity, and stage was confirmed.

A study carried out to provide a biological assessment using an ant-based multimetric index (MMI) sensitive to anthropogenic disturbances in riparian systems, for monitoring, conservation and restoration purposes, also to clarify the extent to which metrics based on ant responses provide useful information that cannot be provided by traditional physical and structural indicators was carried out by [8]. The proposed study is divided into four steps, identify the river typology of the sampling sites, assess the pressure gradient for each river type, develop the ant-based MMI and compare the results obtained by the new ant-based index with those of a traditional physical and structural index. The results presented 2268 individuals comprising 22 ant species, 13 genera, and four subfamilies were identified in the study area. With a collinearity analysis, six metrics showed significant differences between disturbed and less disturbed in the upland river type, while three

metrics allowed the separation between disturbed and less disturbed in the lowland river type. For the upland river type, five main metrics were used: observed species richness, closed habitat species, large ants, cryptic ants, opportunistic ants, and *T. caespitum* foraging activity. For the lowland river type, foraging activity of ants, seed foragers, and *A. senilis* were used.

Considering that ants are easy to measure, easy to sample, sensitive to environmental stress, important in ecosystem functioning, wide distribution, high abundance and predictable responses to environmental stress, in [9] these insects were used as bioindicators to assess ecosystem health of the Northern-Indian Shivalik mountain range, through measurement of diversity, community patterns, species composition and the influence of invasive species of Formicidae; the studies were conducted in 75 sites from 44 locations in three habitats: primary forest (PF), secondary forest (SF), and nonforest areas (NF) using six collection techniques. The research obtained the most comprehensive dataset compiled for Indian ants to date (sample coverage 94% to 97%) and sampled 31,487 ant specimens, representing 181 species from 59 genera and 9 ant families.

In [10], research is conducted to explore benefits and limitations of applying a self-organization approach to ant social structures; it shows how self-organization captures some of the diversity of ant social patterns; and it studies the spatiotemporal structures in ant societies, considering an analogy with self-organization patterns observed in physical and biochemical systems. The research chooses the myrmecological pattern of foraging to examine several properties that are typical features of self-organization. It shows how fluctuations, colony size, and environmental parameters influence the dynamics of feedback loops between interacting ants and thus shape the collective response of the entire ant society. It highlights the capacity of each ant to process information and how ants adjust their interactions with their nestmates, as well as the functional properties of social patterns. The results obtained show that self-organization is a powerful set of pattern-generating mechanisms, or, in other words, a potent generator of biological diversity. Self-organization provides an answer to a challenge in evolutionary biology: how can we generate a wide variety of social patterns in group-living animals while keeping the number of behavioral rules and physiological attributes that should be "encoded" in individuals limited.

[11] uses ecologically relevant behavioral traits to assess colony-level behavioral syndrome in rock ants. Using field and laboratory assays, they measure foraging effort, colony responses to different types of resources, activity level, threat response, and aggression. In their research, the authors find a colony-level syndrome that suggests that colonies differ consistently in their coping style: some are more risk-prone, while others are more risk-averse. Also, by collecting data across their study areas for this species in North America, they find that environmental variation can affect how different populations maintain consistent variation in colony behavior.

An extensive study of ant-plant interactions is presented in [12]; this work investigates ant behavior controlled by plant volatile organic compounds (VOCs); through in situ experiments, it is shown that pollen-derived volatiles can specifically and transiently deter ants during dehiscence; this study proposes the existence of three main types of adaptations that many plant species use to protect their valuable floral structures from ants: architectural barriers; decoys and bribes (e.g., food or lodging located some distance from the flowers); and chemical deterrents, often using floral volatile organic compounds (VOCs).

In [30] the various communication methods used by ants to direct their nest mates to feeding sites, as well as their recruitment behavior for other foraging ants through dances or direct physical contact at the nest, are studied; a broad analysis of the quantities and types of pheromones that ants emit while searching for food is described; also the communication behaviors and interactions that ants perform through pheromone secretion. The functions that each individual in the colony performs are broadly established in this research.

Other research works have carried out studies of ant behavior using mathematical tools, for example in [31] a process algebra approach to modeling ant colony behavior is described; the objective of the work is to show how process algebras can be usefully applied to understand the social biology of

insects, and to study the relationship between the algorithmic behavior of each insect and the dynamic behavior of its colony.

### 1.2. Bees

A wide range of studies on honeybee behavior and its use as bioindicators can be found in the literature. In this subsection, we summarize some of the work related to these two topics. The focus of each research project and the results obtained in each area are described.

A comprehensive set of experimental protocols for detailed studies of all aspects of honey bee behavior is presented in [13]. In this research, honey bees are prepared for behavioral assays to quantify sensory response to gustatory, visual, and olfactory stimuli, appetitive and aversive learning under controlled laboratory conditions, and free-flight learning paradigms; this is done to investigate a wide range of cognitive abilities in honey bees. To explore changes in temperature, experiments analyzing honey bee locomotion across temperature gradients are presented and described in this study.

A work aimed at precision beekeeping that analyzes the development of integrated monitoring systems with multiple sensors and automated motion tracking devices to connect information on bee behavior, hive functioning, and the external environment, in order to explore new fields in bee biology, is presented in [32]. Within this study, the parameters and technologies for obtaining long-term data on the behavior and ecology of bees inside and outside the hive are listed. Through these systems, the behaviors studied include: in-nest behavior and foraging activity, foraging activity and spatial behavior, and spatial behavior and social interactions.

Other works address behavioral aspects of bees, such as in [33] the evolutionary history of eusociality and the antiquity of eusocial behavior in apid bees are studied, as well as swarming behavior (characteristics of highly eusocial colonies), the organization and genetic basis of social behavior in honey bees and their relatives. In [29], an animal model was developed to study how the immune and nervous systems interact in a coordinated manner during an infection in honey bees; the study is carried out by administering a bacterial lipopolysaccharide (LPS) into the thorax of honey bees; this research arises from the assumption that animals undergo various changes in their normal physiology and behavior, which may include lethargy, loss of appetite, and reduced grooming and general movement during the presence of a disease.

A system for automatically analyzing and labeling honeybee movements by creating a behavioral model from examples provided by an expert is presented in [3]. Starting from a manually labeled training data set, the system automatically completes the entire behavioral modeling and labeling process. Experiments involve video recording activity in a hive with a vision-based tracker; this information is considered the location of each animal, and then numerical features such as speed and direction change are extracted. By combining kernel regression classification techniques and hidden Markov models (HMMs), the behaviors generated by the movements are recognized, the sequence of movements of each observed animal is labeled, and evaluations are performed with hundreds of honeybee trajectories.

Considering honeybees as eusocial insects, [34] studies how gene regulatory network (GRN) activity influences developmental plasticity and behavioral performance at the individual level among worker bees. The proposed study tests the hypothesis that individual differences in behavior are associated with changes in brain GRN activity (i.e., changes in the expression of transcription factors and their target genes). Making use of automatic behavior tracking, genomics, and the broad behavioral plasticity present in honeybee colonies with laying workers, this research presents results with key insights into the mechanisms underlying the regulation of individual differences in behavior by brain GRNs.

### 1.3. Scene Graph Generation

This section provides a summary of some articles related to scene graph generation, which have served as the basis for this research work.

We start from two comprehensive surveys of scene graphs (consulted at [1,35]), which contain extensive references and analyses of published research papers. Current methods for generating images based on scene graphs were found in [36]. Evaluation metrics, examples of deep learning benchmark datasets, and knowledge graph-based approaches were found in [37]. In [38], a set of SGG applications is listed, such as image retrieval, visual reasoning, visual question answering, image captioning, structured image generation, and robotics. We have found GGS models in [39], where a relational embedding module is proposed that enables a model to jointly represent connections between all related objects, rather than focusing on a single object, which benefits the classification of relationships between objects. In [40] a framework for semantic image retrieval is developed with two abstractions: a scene graph to describe the scene, and a scene graph base that associates the scene graph with the image.

Finally, in [41], less-explored areas are mentioned, such as the detection of distant object relationships and the detection of social relationships between humans to infer social ties, in addition to extracting hidden social interactions from visual data.

## 2. Project Objectives

### 2.1. General Objective of the Project

To design a distributed computing architecture with IoT devices and a parallel computing architecture to acquire and process images and videos of insects in their natural environment to recognize, classify, and analyze their living behavior at nest entrances, for the sake of preserving and protecting species that are beneficial to humans and the environment.

### 2.2. Specific Objectives

1. Construction of the visual dataset “Bees” consisting of 100 images with 15 objects, 12 attributes and 15 relationships between pairs of objects, to allow the modeling of the behavioral relationships of bees at the nest entrance.
2. Construction of the visual dataset “Ants” consisting of 100 images with 15 objects, 12 attributes and 15 relationships between pairs of objects, to allow modeling of the behavioral relationships of ants at the nest entrance.
3. Build three computational prototypes to observe groups of insects in their nests in three different geographical positions (sites), and determine whether three climatic conditions (temperature, humidity and wind speed) are determining factors in the behavior of the observed insects.
4. Develop image descriptions, object descriptions, object attributes, relationships between objects, and question and answer pairs.
5. Recognize the complex shapes of insects to identify the type of living being observed, through automatic recognition of artificial vision shapes.
6. Observe insect behavior at the nest entrance to recognize and classify distinctive life history traits (e.g., behavioral dominance, main food resources, daily activity rhythm) and consequently show distinctive responses (e.g., abundance, species richness) to natural and anthropogenic disturbances.
7. Use automated computing resources to observe and analyze insects and prevent the mass killing of insects, while helping to preserve and care for species that are beneficial to humans and the environment.

## 3. Justification of the Research Project

Observing and recognizing the life forms, behavior, and organizational structures of insects such as bees and ants is possible through the use of mechanical, electrical, and electronic devices operated by computer software. This is done to prevent collateral damage to the population, the elimination of individuals representing the colony, and mass murder. Therefore, developing a computer system that automatically recognizes and classifies objects in an image, understands the relationships between

objects, and produces a textual description based on the image content is justified for the sake of species preservation.

## 4. Preliminaries

This section presents a set of definitions, which are used in the following sections; subsection 4.1 presents the terminology related to scene graphs, and in the 4.2 subsection, the terminology for scene graph construction is presented.

### 4.1. Scene Graph Terminology

Formally, a graph  $G$  contains a set of nodes  $V$  and a set of edges  $E$ ; each edge  $e_{i \rightarrow j} \in C_{rel}$  defines a relation predicate between the subject node  $V_i$  and the object node  $V_j$ .  $C_{obj}$  is a set of object classes,  $C_{rel}$  is a set of relations [39].

The Scene Graph (SG) is a direct graph data structure, and is defined as:

$$SG = (O, R, E) \quad (1)$$

Where,  $O = \{o_1, \dots, o_n\}$  represent the set of objects detected in the image and  $n$  is the number of objects. Every object can be denoted as  $o_i = (c_i, A_i)$ ,  $c_i$  represent the category of the object and  $A_i$  are the attributes of the object.

$R$  are relationships between nodes, i.e., relationships between the  $i$ -th and  $j$ -th object instance is expressed as  $r_{i \rightarrow j}, i, j \in \{1, 2, \dots, n\}$ .

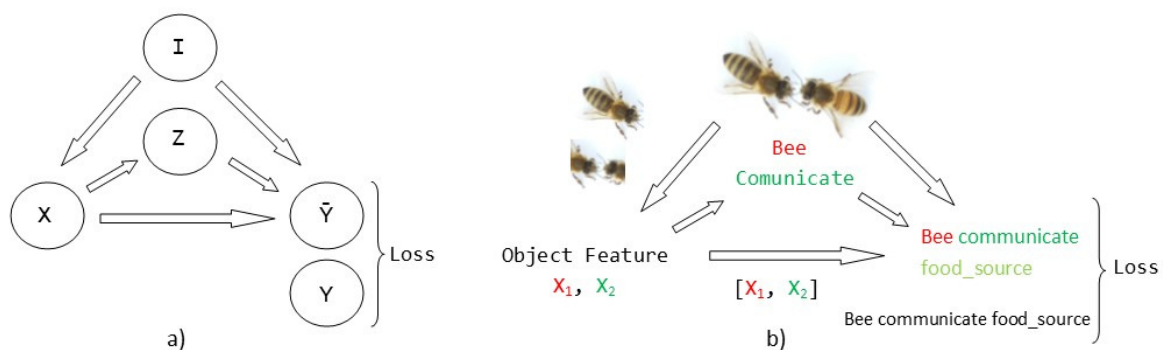
$E$  represents the edges between the object instance nodes and relationship nodes; so there are at most  $n \times n$  edges in the initial graph; then  $E \subseteq O \times R \times O$ .

Considering the representations of scene graph,  $I$  represents a given image, and the SGG produces as output a scene graph with the instances of the objects found in the image, represented with the bounding boxes and the relationships that may exist between the objects. According to [1], we represent the above as:

$$SG_{O,R,E}^I = SGG(I) \quad (2)$$

### 4.2. Scene Graph Construction

To define the process of building a scene graph, consider Figure 1. Point a) in Figure 1 is a graph with 5 nodes and 5 edges (as the example of [1]),  $I$  represents an input image,  $X$  represents the feature of the object and node  $Z$  represents the category of the object.  $Y$  represents the category of the prediction predicate and the corresponding triple  $\langle \text{subject} - \text{predicate} - \text{object} \rangle$ . A fusion function acts as a receiver of the outputs of the three branches and generates the final quantile function, associated with the standard logistic distribution. Point b), in Figure 1 exemplifies construction of a scene graph for a behavior exhibited by bees at the entrance of the hive, using the graph from point a) of this same figure.



**Figure 1.** Process of building a scene graph and its corresponding application.

**Object Feature Extraction** represented by the link  $I \rightarrow X$ . Using the technique proposed in [1], a set of bounding boxes are extracted with their corresponding features, then,  $B = b_i | i = 1, \dots, m$ , represents the set de bounding boxes, and  $X = x_i | i = 1, \dots, m$  represents the corresponding features of the input image  $I$ . The process of obtaining the characteristics is represented as:

$$Input : \{I\} \Rightarrow Output : \{x_i | i = 1, \dots, m\} \quad (3)$$

This process iterates to obtain the encoded visual context of each object in the image. In Figure 1, item b) two bounding boxes are extracted and presented in the two left-hand images: the isolated bee and the two bee heads.

**Object Classification** represented by the link  $X \rightarrow Z$ . This process is expressed in the form:

$$Input : \{x_i\} \Rightarrow Output : \{z_i, z_i \in O\}, i = 1, \dots, m \quad (4)$$

In example case, we classify the image as a "Bee Communicate" in the image (greeting between bees), after object feature extraction was construct; the 6.2 section shows this example in more detail.

**Object Class Input for SGG** represented by the link  $Z \rightarrow \bar{Y}$ . Here the paired object label is used  $(z_i, z_j)$ , and a predicate  $\bar{y}_{ij}$  between a pair of objects is predicted by a level  $M$ , called combined embedding layer. The process can be expressed as:

$$Input : \{(z_i, z_j)\} \Rightarrow Output : \{\bar{y}_{ij}\}, i \neq j; i, j = 1, \dots, m. \quad (5)$$

A priori knowledge and a priori statistics are computed in this step [1]. So, for our example, with the a priori knowledge of the image and established in the region description (defined in the 6.2 section), a query is performed to retrieve images depicting scenes similar to the one described by the graph. The agreement between the retrieved images and the unannotated image (scenario [40]) is measured; to determine the agreement, a conditional random field is constructed that models the distribution of all possible relations; an inference set as maximum posterior finds the most likely agreement; the probability of the maximum a posteriori inference is taken as the score that measures the agreement between the retrieved images and the unannotated image.

**Object Feature Input for SGG** represented by the link  $X \rightarrow \bar{Y}$ . The object features, represented by means of paired  $[x_i, x_j]$  allows the prediction of the corresponding predicate. The process is represented as:

$$Input : \{[x_i, x_j]\} \Rightarrow Output : \{\bar{y}_{ij}\}, i \neq j; i, j = 1, \dots, m. \quad (6)$$

This link uncovers the entire context of the information. For the proposed example, we consider all the information obtained by applying all the relationships obtained with the conditional random field as the object feature input for SGG.

**Visual Context Input for SGG** represented by the link  $I \rightarrow \bar{Y}$ . The visual context feature  $v_{ij} = Conv_{s}(RoIAlign(I, b_i) \cup b_j)$  [1] of the joint region  $b_i \cup b_j$  is extracted in this link and predicts the corresponding triplet. This process is represented as:

$$Input : \{[x_{ij}]\} \Rightarrow Output : \{\bar{y}_{ij}\}, i \neq j; i, j = 1, \dots, m. \quad (7)$$

In the example in Figure 1 item b), the corresponding triplet is: food\_source, which constructs the predicate as "Bee Communicate food\_source".

**Training Loss.** Most of the proposed SGG models are trained using conventional cross-entropy losses for the object label and the predicate label; this training is used in this work. To prevent a single link from spontaneously dominating the generation of logits  $\bar{Y}$ , [15] auxiliary cross-entropy losses are added to predict individually  $\bar{Y}$  from each branch.

## 5. Formal Definition of the Research Problem

Considering the above definitions, formal definition of the research problem is defined as:

Given an image  $I$  containing numbered objects of the form  $o_1, o_2, \dots, o_n$  which belong to the set  $\theta$  of known objects and are represented by numbered graphs  $G$  of the form  $G_1, G_2, \dots, G_n$ . The goal is to find  $n$  hidden relationships between the objects of the form:

$$\begin{array}{cccccc} o_1 & \xrightarrow{R} & o_2, & o_3, & \dots & o_n \\ o_2 & \xrightarrow{R} & o_1, & o_3, & \dots & o_n \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ o_n & \xrightarrow{R} & o_1, & o_3, & \dots & o_{n-1} \end{array}$$

To textually describe the image  $I$ , using formal language. Then, the following objectives are met:

- Statement 1 ( $S_1$ ), object recognition. Given  $I$ , obtain  $o_1, o_2, \dots, o_n$  where  $o_1, o_2, \dots, o_n \in \theta$  otherwise any  $o_n \notin \theta$
- Statement 2 ( $S_2$ ), object relationships. If  $S_1$  is true, then for each  $o_1, o_2, \dots, o_n$  create a graph  $G$ , numbered  $G_1, G_2, \dots, G_n$  and creates the bounding boxes contained in the image. Find a relationship  $R_1, R_2, \dots, R_n$  between:

$$\begin{array}{cccccc} o_1 & \xrightarrow{R_1} & o_2, & or & o_3, & or & \dots & o_n \\ o_2 & \xrightarrow{R_2} & o_1, & or & o_3, & or & \dots & o_n \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ o_n & \xrightarrow{R_n} & o_1, & or & o_3, & or & \dots & o_{n-1} \end{array}$$

- Statement 3 ( $S_3$ ), textual description. If, and only if,  $S_1$  and  $S_2$  are true, then the process of building a scene graph is realized and a formal description of  $I$  is generated.

## 6. Materials and Methods

In this section, all the electrical, electronic, and mechanical materials used in the experiments are described. The methodological approach for constructing the bee image dataset and the ant dataset is then described in detail. Likewise, the way in which scene graph generation is produced and applied is described.

### 6.1. Materials

The materials used in this research are classified as materials related to the IoT architecture of distributed computing, materials for generating solar energy at the insect observation sites, the hardware for the parallel computing architecture, and the hardware for the aerial device used. Each of these materials is explained in the following sections.

The distributed computing IoT architecture, called observation prototypes, consists of four sets of devices used to acquire images from the arthropod nests. Each set of devices contains the following elements:

1. Three Raspberry Pi cards, each with 16 GB of memory and an Intel processor, with WiFi connectivity. Technical information: Raspberry Pi 4 Computer Model B Version 8 GB RAM (SonyUK-TechnologyCentre, Wales, UK).
2. Three 4.5 Mega Pixel Raspberry Pi high definition cameras. Technical information: Raspberry Pi High-Quality Camera 12.3 MP7.9mm diagonal image size.

3. Three Arduino UNOs, connected to a Raspberry Pi board, which contains two connected sensors, one for humidity and one for temperature. Technical information: Smart Projects, Ivrea, Italy, based on the ATmega328P. Humidity sensor DHT11. Temperature sensor LM35.
4. Three Protoboard Breadboard 400 points for Arduino.
5. Three Arduino DHT22 temperature and humidity sensors. Power supply 3.3-6V DC, with Operating range humidity 0-100%RH; temperature -40 to 80 Degrees Celsius.

From each of the four sets of devices, a device is extracted and an observation module is built consisting of a Raspberry Pi card, a high-definition camera, an Arduino UNO, a breadboard, and a temperature and humidity sensor; the image Figure 2 shows an observation prototype already built. Each prototype is installed in one of the observation sites, which operate at the same time to obtain comparative data in real time on the behavior of insects in three different sites, where the nests are located. By installing the three observation modules, we can obtain visual information and data in a distributed and parallel manner.



**Figure 2.** Observation Prototype.

To generate the energy needed to power each of the observation modules, the following materials are required:

1. Three solar panels, with the following technical information: Smart Projects, Ivrea, Italy 50 W, 12 Vdc, Polycrystalline, 36 Grade A Cells connected to a voltage regulator and an inverter;
2. A Snc 20 Solar Controller 12/24v Display Kit Plant Lth Cale.
3. A 750W Power Inverter with USB Ports, DC and AC Inverter. Product brand: Truper.
4. A Gray Pneumatic Tubular Wheelbarrow 5.5ft3 Pretul 20646, which is the mobile container to transport the solar energy generator.

The mobile container is transported to the observation sites by the researcher using brute force.

With regard to parallel computing architecture, it consists of the following devices installed in the Data Processing Center (CPD).

1. A server DELL EMC Intel(R) Xeon(R) Silver 4210R CPU 2.40 GHz 20 Cores, with the host operating system LINUX Ubuntu Server 24.04 LTS.
2. A Server Power Edge R330 Intel(R) Xeon(R) CPU E3-1220 V5 3.00 GHz, with the host operating system LINUX Ubuntu Server 24.04 LTS.
3. A Switch Cisco Gigabit 10/100/1000 for server interconnectivity.
4. An uninterruptible power supply, which supplies power to the servers and guarantees the power supply in the event of a power failure.

In addition to the above, a DJI MINI 4 PRO aerial device (drone) is used to fly over the study area. This device is used to acquire the positions of the anthills to be observed. The aircraft flies within 180 hectares of University land.

## 6.2. Method

Using the proposed hardware, a methodological approach was developed, consisting of five steps. The steps are listed below, and each is explained in the following paragraphs.

1. Establish the study area to carry out experiments with honeybees and ants.
2. Creating the dataset.
3. Scene graph generation.
4. Training the neural network.





#### 6.2.1. Establish the Study Area to Carry Out Experiments with Honeybees and Ants

To collect visual information on the bees, there are two test apiaries at the University, which are used in this research project. To acquire information from other apiaries, the image acquisition equipment is brought to the site.

To locate the anthill locations and establish access points to the nests, a drone equipped with a video camera and transmitted to a cell phone was used. The videos acquired during the overflights are analyzed to detect the anthills' geographical locations. Once the anthills are located, the image acquisition team moves to the positions found to collect the visual information.

Table 1 shows images of four examples of geographical locations and aerial photos of the four sampling sites at the university. The red circles show the exact location of the anthills in the images. The sampled sites correspond to different lands; for example, image 1 corresponds to an anthill between a corn crop area and a bean crop area. Image 2 shows an anthill near a tree and next to pig and sheep farms. Image 3 is an anthill located in an area with cactus plants, where ants do not have access to the crops. Image 4 shows an anthill near a building and walkways for people.

**Table 1.** Geographical locations and aerial photos of four sampling sites of anthills.







Image 1. Anthill 1	Image 2. Anthill 2
	
Image 3. Anthill 3	Image 5. Anthill 4
	

#### 6.2.2. Creating the Dataset

There are methods proposed in the literature for the creation of datasets [1,42]. In this work, the following steps for the generation of the datasets and SGG were considered: images acquired at the observation sites, the description of regions, objects, attributes, relationships, the generation of region graphs, the generation of scene graphs, and the question-answer pairs. The way in which each of these components is applied to the construction of the "Bee" and "Ants" datasets is described in the following paragraphs.

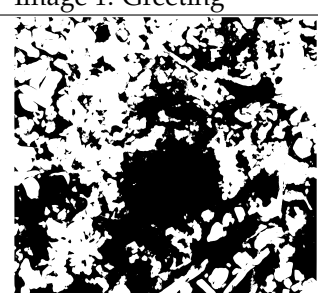

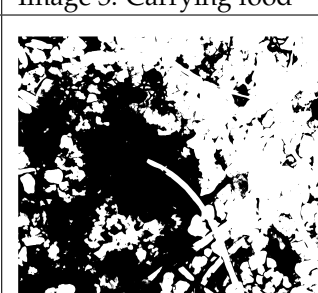
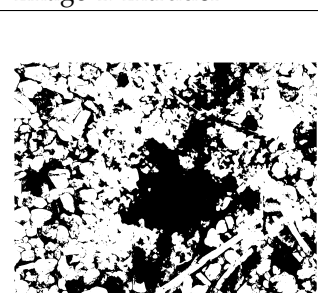
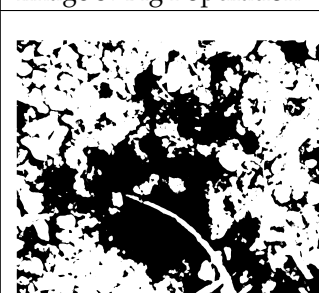
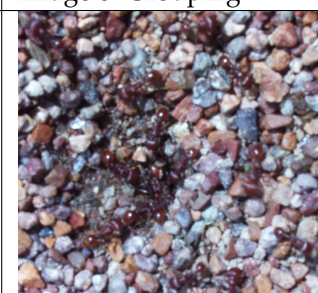
*Images acquired at the observation sites.* Table 2 contains examples of images obtained from access to hives in different apiaries, and Table 3 are images of entrances to different anthills. To more easily identify each image in the tables in the following explanations, each image is assigned a number from left to right starting from the top left corner; in addition to the number, each image is assigned a name related to the behavior that the bees show. Thereby, in Table 2 Image 1, labeled "Greeting" shows two bees greeting each other; Image 2, "Bee Type" shows two types of bees: the Italian honey bee and the Africanized honey bee, distinguished by their yellow and black abdomen; Image 3, "Carrying Pollen" shows a bee carrying pollen on its limbs. Image 4, "Intruder" shows an intruder (a fly) at the hive entrance. Image 5, "Big Population" shows a subset of bees on the hive access platform. Image 6, "Grouping" shows a group of four bees transmitting information about potential food sources.

Table 2. Examples of images acquired from the test apiaries.

Image 1. Greeting	Image 2. Type of bee	Image 3. Carrying food
		
Image 4. Intruder	Image 5. Big Population	Image 6. Grouping
		

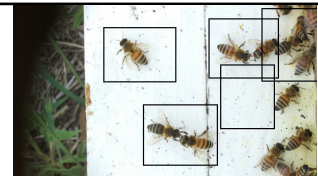

Now, for the case of the ants in Table 3, they have been assigned a number from left to right starting from the top left corner in the same way; the name assigned to each of the images corresponds to the behavior that the ants show.

Table 3. Examples of images acquired from the anthills.

Image 1. Greeting	Image 2. Type of ant	Image 3. Carrying food
		
Image 4. Intruder	Image 5. Big Population	Image 6. Grouping
		

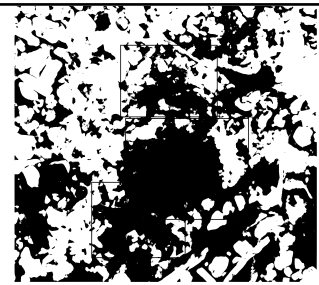
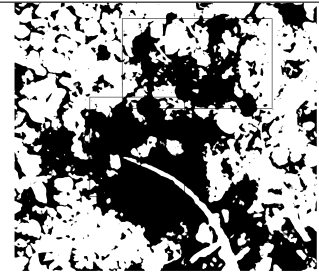
*Region Descriptions.* Explaining or generating the description of an image is called image captioning or frame captioning [37]. The descriptions of the scene regions were made by experts and located by bounding boxes; overlaps between regions are allowed when the descriptions differ; each sentence describing each of the regions varies from 1 to 16 words in length. Five descriptions of regions per image were considered. Table 4 illustrates the descriptions of two images that appear in Table 2: image 1, “Greeting”, and image 4, “Intruder”. For the sake of simplicity, in the rest of this document, we use the word “access” to refer to the beehive entrance or anthill entrance.

**Table 4.** Descriptions of the regions of two scenes from the Table 2: Greeting and Intruder.

	Bee Greeting Beehive Entrance Bee Take-Off Platform Isolated Bee at Beehive entrance Bees Loitering
	Intruder detected Beehive exit Carrying pollen Bee flying Bee lateral position

The descriptions of the scene regions, in the case of the ants, are considered image 1, “Greeting”, and image 5, “Big Population”. Table 5 illustrates the descriptions of two images that appear in Table 3. Because ants have more limited interactions at the nest entrance, in this research only 5 descriptions of the regions of two scenes are considered.

**Table 5.** Descriptions of the regions of two scenes from the Table 3: Greeting and Big Population.

	Ant Greeting Anthill Entrance Isolated Ant at Anthill
	Carrying Food Ants Loitering

*Multiple objects and their bounding boxes.* Once the bounding boxes have been drawn on the images, an average of 5 objects are detected within them; each object is bounded by a tight bounding box and each object is canonicalized to a synset identifier in WordNet [43]. For example, bees would map to bee.n.03 (which is the generic use of bee to refer to a bee), mellifera bee maps to bee.n.01 (a bee). These two concepts can be joined to bee.n.01 since it is a hypernym of bee.n.03. Standardization is achieved with the WordNet ontology for naming an object with multiple names (e.g., bee, apis mellifera, apis mellifera bee) and for connecting information between images. Table 6 shows the mapping from sentences to bounding boxes for the two images in Table 1; from item a) to item e) each sentence is mapped to describe each of the regions of the bounding boxes.

**Table 6.** Mapping the phrases to the bounding boxes of the two images in Table 1.

	<ul style="list-style-type: none"> <li>a) Bee greeting</li> <li>b) Beehive Access</li> <li>c) Bee take-off platform</li> <li>d) Isolated bee at the entrance</li> <li>e) Bees Loitering</li> </ul>
	<ul style="list-style-type: none"> <li>a) Intruder detected</li> <li>b) Beehive exit</li> <li>c) Carrying pollen</li> <li>d) Bee flying</li> <li>e) Bee lateral position</li> </ul>

The same is true for ants, the bounding boxes are drawn on the images, with an average of three objects; each object is bounded by a tight bounding box and each object is canonicalized to a synset identifier. For example, ant would map to ant.n.03 (which is the generic use of ant to refer to an ant), emmet bee maps to ant.n.01 (an ant), since it is a synonym of ant.n.03. Just like the terms used to refer to a bee, we refer to an ant by standardizing with the WordNet ontology for naming an object with multiple names, for connecting information between images. Table 7 shows the mapping from sentences to bounding boxes for the two images in Table 3; from item a) to item e) each sentence is mapped to describe each of the regions of the bounding boxes.

**Table 7.** Mapping the phrases to the bounding boxes of the two images of ants in Table 3.

	<ul style="list-style-type: none"> <li>a) Isolated ant at the entrance</li> <li>b) Anthill Access</li> <li>c) Ant greeting</li> </ul>
	<ul style="list-style-type: none"> <li>a) Carrying Food</li> <li>b) Ants Loitering</li> </ul>

*A set of attributes.* Once the objects are collected from the region descriptions, the attributes associated with these objects are also collected. The attributes are canonized in WordNet [43]. Table 8 shows three attributes: color, state and position, associated with the objects: Bee, Intruder and Beehive Access.

**Table 8.** Objects and attributes associated with objects (Bees).

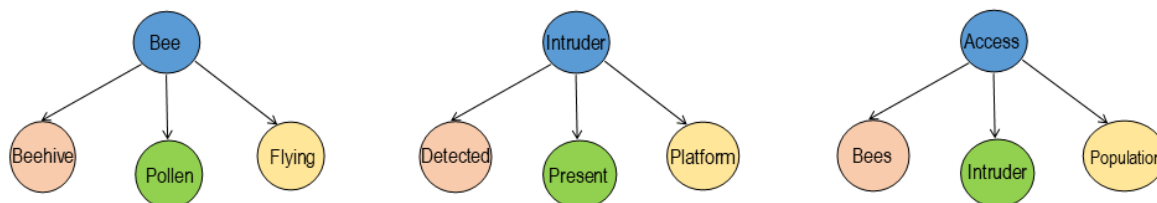
Object	Attributes associated with objects		
	Color	State	Position
Bee	Yellow	Flying	Raised wings
	Black	Accessing	Wings at rest
	Grey	Leaving	Next to
Intruder	Grey	Vigilant	Raised wings
Access	Free	Without bees	Vertical
	Occupied	Low population	Vertical
	Tumult	High population	Vertical

Now, in the case of ants, the selected attributes have been matched to the types of ants observed; in our case, the colors of the ants are red, yellow, and black. Although in the case of intruders, it is very difficult for intruders to approach the access in an anthill because ants display very aggressive behavior and are very jealous of their spaces, within the dataset we have recorded the arrival of intruders to the accesses of the anthills. In the case of access to the anthill, the attributes are similar to those of bees; the access can be Free, Occupied, or Tumult with their respective associated attributes. Table 9 shows objects and attributes associated with objects (Ants).

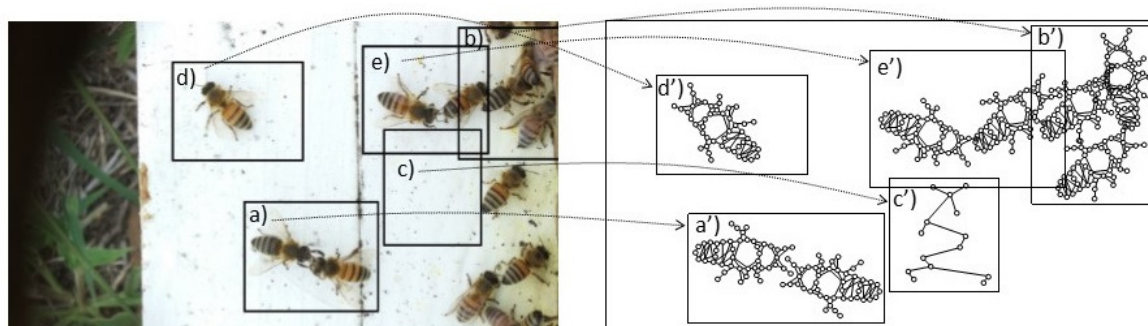
**Table 9.** Objects and attributes associated with objects (Ants).

Object	Attributes associated with objects		
	Color	State	Position
Ant	Red	Walking	Raising antennas
	Yellow	Accessing	Detained and guarded
	Black	Leaving	Next to
Intruder	Grey	Vigilant	Raised wings
Access	Free	Without ants	Vertical
	Occupied	Low population	Vertical
	Tumult	High population	Vertical

*Set of relationships.* Relationships connect objects, with descriptive verbs, propositions, actions, comparisons, or propositional phrases. A relationship is directed from one object to another, that is, from the subject to the object. Relationships are canonicalized to an identifier from a WordNet synonym set [43]. Figure 3 shows three examples of relationships that connect the subject to the object. In example 1, the object Bee is related to three objects: the object Beehive, the object Pollen, and the object Flying.

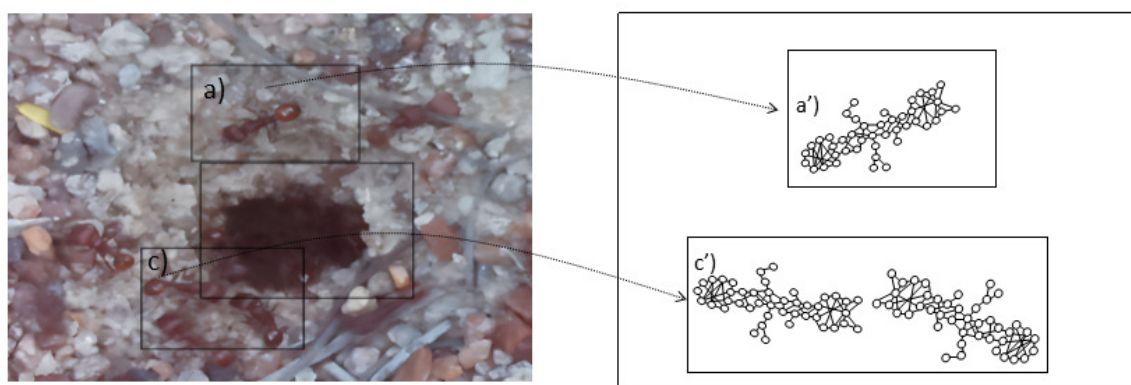
**Figure 3.** Examples of relationships which connect subject - object.

*Set of region graphs.* In this stage, a graph representation that addresses each of the regions is created; a region graph represents a structured part of the image. Objects, attributes, and relationships are represented by the nodes of the graphs. Objects are linked to their attributes, and relationships link one object to another. Figure 4 shows an example of the set of region graphs produced for image 1, namely, Greeting from Table 2.



**Figure 4.** Example of a set of region graphs produced for image 1 in Table 2.

The set of region graphs of ants is also shown in the Figure 5; The way in which the Isolated graphs are created at the nest entrance and the ant greeting, which were shown in Table 3, are also exemplified in this figure.

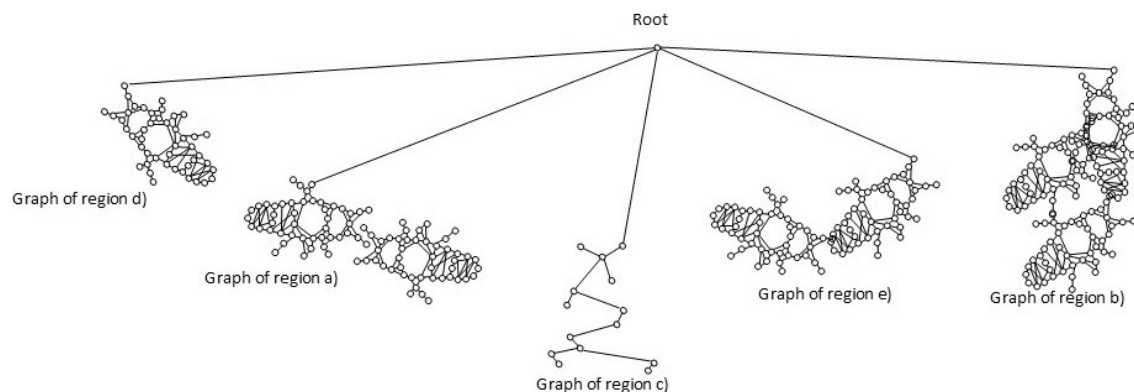


**Figure 5.** Example of a set of region graphs produced for image 1 in Table 3.

### 6.2.3. Scene Graph Generation

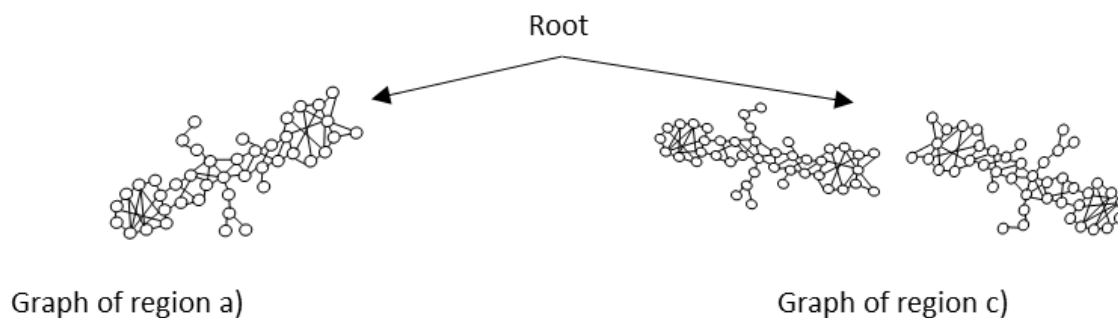
*Scene Graph.* Of the existing methods for scene graph generation in the literature [1,35], in this work we have used the Graph Neural Network (GNN) [44–46] based method for the following justifications, first, in [27,28] a procedure for converting the entire image to a graph is described; by using the described method, in this research a set of graphs are constructed in the step of generating the set of region graphs (described in the previous paragraphs), where each region is considered a graph. Then, by the above, the region graphs are combined into a single scene graph that represents the entire image, that is, a scene graph is the union of all the region graphs and contains all the objects, attributes and relations of the description of each region, which allows combining multiple levels of scene information in a more coherent way. Section 6.2.4 describes how the neural network is trained, including the input to the neural network, as well as the embedding vectors and the generation of the network's outputs.

Figure 6 is an example of the set of region graphs produced for image 1 in Table 2, combined into a single scene graph. Once the scene graph has been generated, graph reasoning is performed to complete the procedure; in this paper, the complete procedure for graph reasoning is described in following paragraphs.



**Figure 6.** Example of a set of region graphs produced for image 1 in Table 2.

Now for the case of the ant scene graph, Figure 7 shows the set of region graphs produced for image 1 in Table 3.



**Figure 7.** Example of a set of region graphs produced for image 1 in Table 3.

*Set of question and answer pairs.* Once the procedure for generating scene graphs from the images is complete, the final component is defining the set of question-answer pairs. To define the set of question-answer pairs, two types of questions and answers are associated with each image in the dataset; the first pair is free-form based on the entire image; the other pair refers to questions and answers based on selected regions of the image. The question types per image are: what, where, how, when, who, and why.

Table 10 shows a set of free-form questions and answers based on the complete image based on the full image. The results obtained with this example are described in section 8 Experiments.

**Table 10.** Set of free-form questions and answers based on the complete image.

Question	Answer
How is the access to the beehive?	Big Population
What are the bees doing at the entrance to the hive?	Grouping
Who is at the entrance of the hive?	Intruder
Why are the bees working?	Carring pollen
What is in the beehive entrance?	Type of bee

#### *Graph Reasoning*

This section describes graph reasoning for sentence detection, predicate classification, and scene graph classification as referenced in [41].

*Phrase detection.* In this phase, an output text known as a label is generated (*subject – predicate – object*). The process for detecting relationships between the subject and the object is located using a bounding box, as defined in section 4 Preliminaries and is referenced in the Table 3.

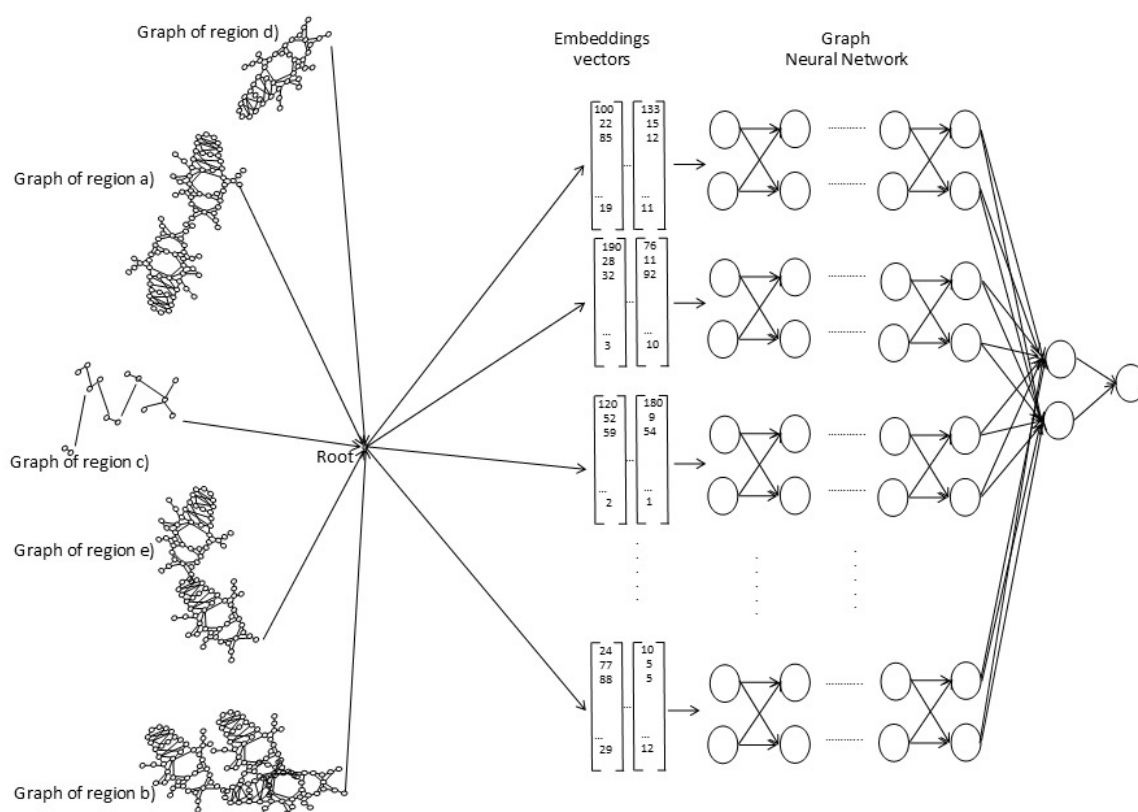
*Classification of predicates.* For this phase, a set of object pairs is generated within the image, multiple objects and their bounding boxes as explained in subsection 6.2.2; from this set of object detected pairs, it is determined whether the pairs interact and the predicate of each pair is categorized.

*Classification of scene graphs.* Scene graphs take as input the located objects, predict the categories of the objects and predict the predicates in each pairwise relationship.

The components of the method have been described so far; the following subsections explain the training of neural network from the scene graph generation and the evaluation method.

#### 6.2.4. Training the Neural Network

The neural network model used is the Graph Neural Network, as described in section 6.2.3. Graph Neural Networks are trained for object recognition and classification in the scene, with the region graphs of images. For each of the region descriptions (shown in Table 4, for bees, and in Table 5, for ants), a dataset is created; the graphs of each image in the dataset are converted into embedding vectors that serve as input to the neural network. The embedding vectors then contain all the features of the image and are converted into useful vector representations. The graph neural network (GNN) computes the vector representations at each vertex of the input graph using a transition function  $F$ , which allows computing the representation of the next neighborhood from the current representation; consider that the neighborhood size is variable, then the transition function is applied symmetrically. Figure 8 is a general representation of the training of the neural network for the recognition, classification and description of objects in the scene.



**Figure 8.** Training of the neural network from the scene graph generation.

The three components of the network are mentioned in the following list.

- Input to the neural network. The Graph Neural Network receives the complete scene graph of the image as input; generated from the bounding boxes; the object characteristics are obtained for the generation of vector embeddings as performed in [28].

- Processing. The Graph Neural Network receives the embedding vectors as input and generates recognition of the objects in the scene.
- Neural network output. The output produced by the neural network is object recognition and classification.

### 6.3. Evaluation of the Method

In [28] method is evaluated in two ways: first, the system interacts with a user, and second, the system runs automatically. When the user interacts with the system, they choose an image and can freely formulate a question. The question formulated is based on the regions of the image or use the set of questions proposed in Table 10. Then, the scene graph is generated and the resulting sentence or predicate is sent to the user. In contrast, when the system runs automatically, the host server receives the images remotely from the apiary, generates the scene graphs, and provides the results.

In this work, we performed the following experiments: first, automatic object recognition by the system for both types of arthropods: bees and ants; second, the generation of bounding boxes from the images and scene recognition with the generated sentences. Each experiment is described in the 8 section along with its results.

## 7. Architecture of Distributed Artificial Intelligence

Before describing the experiments carried out, this section presents the architecture of the distributed artificial intelligence designed and used in this research.

Numerous literature studies indicate that arthropods are susceptible to climate change. However, observing only a single nest provides limited information on the current status of a single colony. Therefore, this research proposes parallel, real-time observations of three geographically distributed sites. We begin with the research question: Is there a direct relationship between the behavior of arthropods (bees and ants) that inhabit colonies in geographically distributed sites with similar ambient temperatures and relative humidity?

This experiment proceeds as follows: a distribution of three prototypes (like the one shown in Figure 2) are divided into three different sites. At each of the three sites, each prototype acquires images of the nest entrances, which will allow us to understand the behavior of the different bee and ant colonies under certain climatic conditions: ambient temperature and relative humidity. We measure the impact on the behavior of each nest. For example, we observe whether the three nests behave the same when the ambient temperature is 14 degrees Celsius and the relative humidity is 42%.

To do this, the prototypes are installed at the observation sites, then the images are acquired in the real environment, transmitted to the data processing center and the massive behavior of the nests is determined using the descriptions presented in Tables 4 and 5.

## 8. Experiments

As metrics are presented in the literature [37] to determine the accuracy of image descriptions, in this work, we have decided to start with the evaluations of the techniques for object recognition, bounding boxes and scene recognition; later, in future works other metrics will be used to measure the accuracy in graph reasoning for sentence detection, predicate classification, and scene graph classification [41]. Then, three parameters evaluated in this research work are, first, the percentage of accuracy in the recognition of object, second the percentage of accuracy in bounding boxes, and finally the percentage of accuracy in scene recognition. The experiments are carried out with the “Bees” and “Ants” datasets separately. This evaluation allows us to determine the effectiveness of the method’s execution when it is executed in real time.

Fifteen datasets with different numbers of images were created for each evaluation; the first dataset contains 100 images, and the last one contains more than 1500 images. The same data sets were used for the evaluation of recognition, bounding boxes and scene recognition.

### 8.1. Object Recognition

To evaluate the percentages of accuracy in the recognition of objects and their bounding boxes, the system receives the image and performs the detection of objects. In this experiment ( $S_1$ ) must be true:

Statement 1 ( $S_1$ ), object recognition. Given  $I$ , obtain  $o_1, o_2, \dots, o_n$  where  $o_1, o_2, \dots, o_n \in \theta$  otherwise any  $o_n \notin \theta$

The GNN recognizes the objects in the image. For each execution of the system, with each of the datasets used, cases in which the program output is correct are counted, as well as the cases in which the program does not output the correct result; the total for each of these cases is divided by the number of images in the data set, to obtain percentages of accuracy in the recognition of objects and their bounding boxes. Figure 9 shows the results obtained from the evaluations performed.



**Figure 9.** Percentages of accuracy in the recognition of objects and their bounding boxes.

*Comments.* According to Figure 9, in these experiments, we observed that as the number of images in each dataset increases, recognition rate remains stable in contiguous data sets, and then decreases. To prevent a decrease in recognition accuracy, as shown in [28], we expanded the range of behaviors of bees and ants that the system needs to analyze. In other words, by incorporating more types of behavior, we were able to improve the accuracy of the recognition system. Additionally, to address the challenges posed by varying positions of bees and ants, the datasets were expanded with more images, and a more comprehensive catalog of images depicting intruders attempting to access the nests was included. However, special care must be taken with overtraining the neural network, because increasing the number of training instances can cause the neural network to become blocked; the percentages reported in this research were obtained with images that represent most of the behaviors of arthropods exemplified in Tables 3 and 4 of section 6.2.2.

### 8.2. Bounding Boxes Evaluation

It is clear that the accuracy of bounding box detection leads to excellent generation of image predicates. Therefore, we consider it a high priority to conduct the experiment to evaluate the accuracy of the bounding boxes generated in the images. For the above, Statement 2 is evaluated as follows:

Statement 2 ( $S_2$ ), object relationships. If  $S_1$  is true, then for each  $o_1, o_2, \dots, o_n$  create a graph  $G$ , numbered  $G_1, G_2, \dots, G_n$  and creates the bounding boxes contained in the image. Find a relationship  $R_1, R_2, \dots, R_n$  between objects.

The bounding boxes evaluated correspond to the 10 frames that can be generated for bees and are shown in Table 6. For ants, the creation of the 5 bounding boxes shown in Table 7 was evaluated.

This experiment is performed as follows: an image dataset containing the bounding boxes was trained; then, the system receives as input the images of the observation sites (established in the 15 datasets represented in the Figure 10 on the X axis), detects the objects in the image, produces the graph and generates the bounding boxes in the received image. To evaluate the accuracy of the output, i.e., the detection of bounding boxes in the image, it is verified against the training image dataset; this verification procedure allows us to determine the accuracy of detecting at least 4 out of 6 bounding boxes in the image. If 4 or 5 bounding boxes were detected, then the experiment is considered correct; if only 3 or fewer bounding boxes are detected, the detection is considered a failure. Figure 10 shows the recognition percentages obtained with 15 datasets.

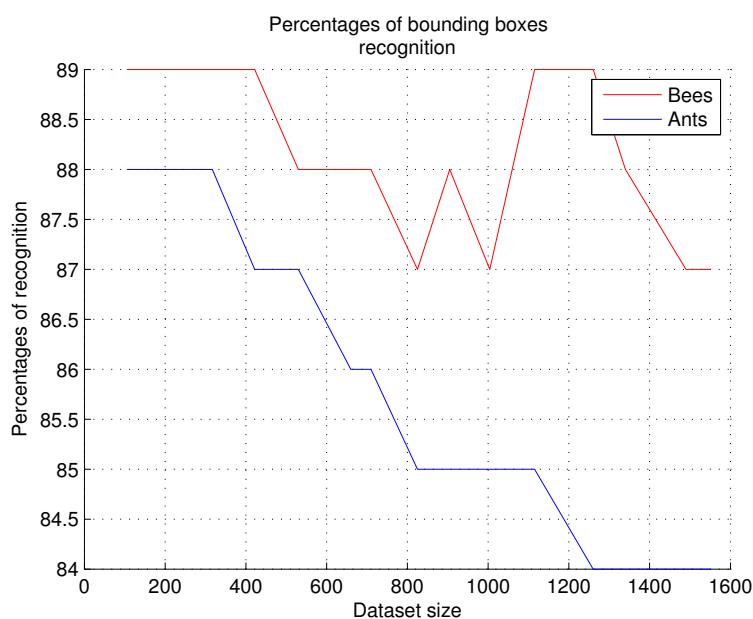


Figure 10. Evaluation of bounding boxes.

*Comments.* According to Figure 10, recognition rates of up to 89% have been obtained in these experiments when the bees' bounding boxes are obtained. In the case of ants, the recognition rates for bounding boxes are lower; the highest rates reach a maximum of 88%. We have observed that only images showing new positions of bees or ants are not recognized by the neural network; for example, when bees fly over the entrance to the nest, and in the case of ants, bounding boxes are not generated when the objects overlap. When new positions of these arthropods are found, they are then included in the training dataset for future recognition.

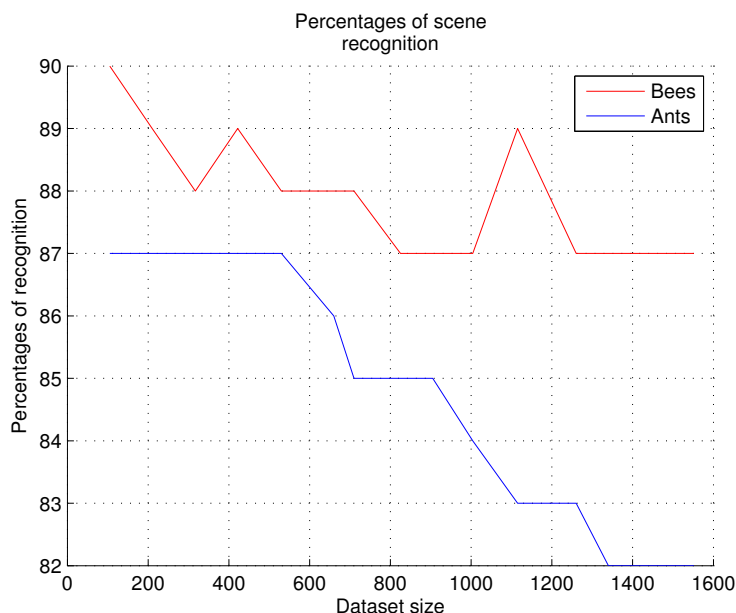
### 8.3. Scene Recognition

Once the two previous experiments have been carried out, we are able to perform Scene Recognition to obtain textual description of the image. Then statement 3 is evaluated:

Statement 3 ( $S_3$ ), textual description. If, and only if,  $S_1$  and  $S_2$  are true, then the process of building a scene graph is realized and a formal description of  $I$  is generated.

In these experiments we evaluated the accuracy percentages in scene recognition. The results were obtained as follows: when the system categorizes the predicate for each object pair, it is determined whether the predicate is correct for the image or incorrect, using Visual Context Input for SGG described in section 4.2. To determine if a predicate is correct we use Training Loss, as defined in the 4.2 section. The predicates evaluated as correct and the predicates evaluated as incorrect are counted for each dataset evaluated; the number of correct predicate evaluations is divided by the number of images

recognition in the dataset to obtain the percentages of accuracy in scene. Figure 11 shows the scene recognition percentages obtained with 15 datasets.



**Figure 11.** Evaluation of scene recognition.

In this experiment, predicate categorization rates for each object pair were identified as high as 90% for the bees and as high as 87% for the ants, as shown in the Figure 11. To ensure 90% recognition, it was necessary to achieve similar percentages in the two previous experiments. We can conclude from this experiment that the 10% non-recognition rate for the bees is due to complex scenes that show overlapping colony individuals or because the nest entrance remains empty. Other factors, such as a cloudy day or shadows cast by the prototype at the nest entrance, are also causes of scene non-recognition. This experiment also allowed us to identify how the triplet  $\langle \text{subject} - \text{predicate} - \text{object} \rangle$ , (we refer to the section 6.2.3), is constructed and to determine whether this triplet is appropriate for the scenes evaluated in each of the images.

## 9. Conclusion

In this research, we conducted an investigation aimed at observing the behavior and organization of two types of insects: bees and ants, using nest accesses as observation areas. A hardware system was developed to obtain visual and environmental information from the research sites, namely, anthills and apiaries. The hardware system was complemented by a software system that uses artificial intelligence techniques, specifically scene graphs, to recognize objects in the scene, determine relationships between objects, and construct a textual description per image. Through this research, we demonstrate that current technology allows us to preserve and care for endangered species, while also preventing the mass killing of insects currently carried out for the purpose of analyzing and studying endemic species. In an extensive literature review, we found no computational system for recognizing behavioral and organizational aspects of bees and ants using the techniques proposed in this research. Extensive experiments were conducted in real-life environments, including access to anthills and beehives. The results of the experiments demonstrate the feasibility of the proposed software system and hardware architecture for recognizing, classifying, and generating textual descriptions of the images obtained.

## 10. Discussion

Referring to the work of [47], Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?, we commented that if that butterfly is eliminated by conducting studies that capture hundreds of insects to be killed and then observed, that tornado will surely not reach Texas. In

the century in which we exist, it is no longer justifiable to eliminate insects to conduct a study; current technologies offer us the opportunity to automate processes, to visualize what our eyes cannot see, but we must develop the systems to do so.

## 11. Patents

A patent for the observation prototype is in process.

**Author Contributions:** Conceptualization, Apolinar Velarde Martinez; methodology, Apolinar Velarde Martinez; software, Apolinar Velarde Martinez; validation, Apolinar Velarde Martinez and Gilberto Gonzalez Rodriguez; formal analysis, Apolinar Velarde Martinez; investigation, Apolinar Velarde Martinez and Gilberto Gonzalez Rodriguez; resources, Apolinar Velarde Martinez and Gilberto Gonzalez Rodriguez; data curation, Apolinar Velarde Martinez; writing—original draft preparation, Apolinar Velarde Martinez; writing—review and editing, Apolinar Velarde Martinez; visualization, Apolinar Velarde Martinez; supervision, Apolinar Velarde Martinez; project administration, Apolinar Velarde Martinez; funding acquisition, Apolinar Velarde Martinez and Gilberto Gonzalez Rodriguez.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** The author is grateful for the administrative support received from Instituto Tecnológico del Llano, Aguascalientes, Mexico and Tecnológico Nacional de México, to carry out all the necessary experiments for this research within their facilities.

**Conflicts of Interest:** Author declare no conflicts of interest.

## References

1. X. Chang, P. Ren, P. Xu, Z. Li, X. Chen and A. Hauptmann. A Comprehensive Survey of Scene Graphs: Generation and Application. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 1-26, 1 Jan. 2023, doi: 10.1109/TPAMI.2021.3137605.
2. Sajjad, Hareem. "Terrestrial insects as bioindicators of environmental pollution: A Review." *UW Journal of Science and Technology* Vol. 4 (2020) Pp. 21-25. University of Wah. Punjab, Pakistán.
3. FELDMAN, Adam; BALCH, Tucker. Representing honey bee behavior for recognition using human trainable models. *Adaptive behavior*, 2004, vol. 12, no 3-4, Pp. 241-250.
4. Anjos, DV., Tena, A., Viana-Junior, A., Carvalho, R., Torezan-Silingardi, H., Del-Claro, K., Perfecto., The effects of ants on pest control: a meta-analysis, *Proc. R. Soc.*, 2022, B.28920221316, 1-11, B 289: 20221316, <http://doi.org/10.1098/rspb.2022.1316>
5. Lutinski, Junir Antonio ; Lutinski, Cladis Juliana ; Serena, Adriely Block ; Busato, Maria Assunta ; Mello Garcia, Flávio Roberto. Ants as Bioindicators of Habitat Conservation in a Conservation Area of the Atlantic Forest Biome. *Sociobiology (Chico, CA)*, 2024-03, Vol.71 (1), p.e9152 Pp. 1-7. DOI: 10.13102/sociobiology.v71i1.9152
6. Wagner D., J. B. Jones, and D. M. Gordon, 2004. Development of harvester ant colonies alters soil chemistry. *Soil Biology and Biochemistry* 36: 797-804.
7. Dominguez-Haydar Y., y I. Armbrrecht. 2011. Response of Ants and Their Seed Removal in Rehabilitation Areas and Forests at El Cerrejón Coal Mine in Colombia. *Restoration Ecology* 19: 178–184.
8. Zina, V., Ordeix, M., Franco, J. C., Ferreira, M. T., & Fernandes, M. R. (2021). Ants as bioindicators of riparian ecological health in catalonian rivers. *Forests*, 12(5), 625.
9. Akhila, A. and Keshamma, E. Recent perspectives on ants as bioindicators: A review. *Journal of Entomology and Zoology Studies* 2022, Vol. 10, issue 3, Pp. 11-14. E-ISSN: 2320-7078. DOI: <https://doi.org/10.22271/j.ento.2022.v10.i3a.9005>
10. DETRAIN, Claire; DENEUBOURG, Jean-Louis. Self-organized structures in a superorganism: do ants "behave" like molecules?. *Physics of life Reviews*, 2006, vol. 3, no 3, p. 162-187.
11. Bengtson, S. E., & Dornhaus, A. (2014). Be meek or be bold? A colony-level behavioural syndrome in ants. *Proceedings of the Royal Society B: Biological Sciences*, vol 281, no.1791, 20140518. p. 20140518
12. WILLMER, Pat G., et al. Floral volatiles controlling ant behaviour. *Functional Ecology*, 2009, vol. 23, no 5, p. 888-900.

13. Scheiner, R., Abramson, C. I., Brodschneider, R., Crailsheim, K., Farina, W. M., Fuchs, S., ... & Thenius, R. (2013). Standard methods for behavioural studies of *Apis mellifera*. *Journal of Apicultural Research*, 52(4), 1-58.
14. Anjos, D. V., Tena, A., Viana-Junior, A. B., Carvalho, R. L., Torezan-Silingardi, H., Del-Claro, K., & Perfecto, I. (2022). The effects of ants on pest control: a meta-analysis. *Proceedings of the Royal Society B*, 289(1981), 20221316.
15. Sampat Ghosh, Chuleui Jung, Victor Benno Meyer-Rochow. Nutritional value and chemical composition of larvae, pupae, and adults of worker honey bee, *Apis mellifera ligustica* as a sustainable food source, *Journal of Asia-Pacific Entomology*, Volume 19, Issue 2, 2016, Pages 487-495, ISSN 1226-8615, <https://doi.org/10.1016/j.aspen.2016.03.008>.
16. Girotti, S., Ghini, S., Ferri, E. et al. Bioindicators and biomonitoring: honeybees and hive products as pollution impact assessment tools for the Mediterranean area. *Euro- Mediterranean Journal for Environmental Integration* 5, 62 (2020). <https://doi.org/10.1007/s41207-020-00204-9>
17. Kadry, A. R. M., Khalifa, A. G., Abd-Elhakam, A. H., Moselhy, W. A., & Abdou, K. A. H. (2025). Honey as a Valuable Bioindicator of Environmental Pollution. *Journal of Veterinary Medical Research*, 32(1), 10-16.
18. ESQUIVEL, Isaac L.; PARYS, Katherine A.; BREWER, Michael J. Pollination by non-*Apis* bees and potential benefits in self-pollinating crops. *Annals of the Entomological Society of America*, 2021, vol. 114, no 2, p. 257-266. doi: 10.1093/aesa/saaa059
19. KHALIFA, Shaden AM, et al. Overview of bee pollination and its economic value for crop production. *Insects*, 2021, vol. 12, no 8, p. 688.
20. Hung Keng-Lou James, Kingston Jennifer M., Albrecht Matthias, Holway David A. and Kohn Joshua R. 2018The worldwide importance of honey bees as pollinators in natural habitats *Proc. R. Soc. B*.28520172140 <http://doi.org/10.1098/rspb.2017.2140>
21. Giovanni Cilia, Laura Bortolotti, Sergio Albertazzi, Severino Ghini, Antonio Nanetti, Honey bee (*Apis mellifera* L.) colonies as bioindicators of environmental SARS-CoV-2 occurrence, *Science of The Total Environment*, Volume 805, 2022, 150327, ISSN 0048-9697, <https://doi.org/10.1016/j.scitotenv.2021.150327>.
22. Método para Determinar Niveles de Varroa en Terreno. Disponible en: <https://teca.apps.fao.org/teca/pt/technologies/8663>.
23. Schurischuster, S., Remeseiro, B., Radeva, P., Kampel, M. (2018). A Preliminary Study of Image Analysis for Parasite Detection on Honey Bees. In: Campilho, A., Karray, F., ter Haar Romeny, B. (eds) *Image Analysis and Recognition. ICIAR 2018. Lecture Notes in Computer Science*, vol 10882. Springer, Cham. [https://doi.org/10.1007/978-3-319-93000-8\\_52](https://doi.org/10.1007/978-3-319-93000-8_52).
24. Chiron, G., Gomez-Krämer, P., & Ménard, M. (2013). Detecting and tracking honeybees in 3D at the beehive entrance using stereo vision. *EURASIP Journal on Image and Video Processing*, 2013, 1-17.
25. Henry J. Svec & Ambarish Ganguly (28 Feb 2025). *Bee World*. Using AI for Hive Entrance Monitoring, *Bee World*, DOI: 10.1080/0005772X.2025.2466903
26. Voudiotis, G.; Moraiti, A.; Kontogiannis, S. Deep Learning Beehive Monitoring System for Early Detection of the Varroa Mite. *Signals* 2022, 3, 506-523. <https://doi.org/10.3390/signals3030030>
27. Martínez, A. V., Rodríguez, G. G., Cabral, J. C. E., & Moreira, J. D. R. (2023, November). Varroa Mite Detection in Honey Bees with Artificial Vision. In *Mexican International Conference on Artificial Intelligence 2023* (pp. 315-330). Cham: Springer Nature Switzerland.
28. Martínez, A.V., Gonzalez Rodriguez, G., Estrada Cabral, J.C. (2026). Recognition of Bee Organizational Behavior with Scene Graphs Generation. In: Martínez-Villaseñor, L., Vázquez, R.A., Ochoa-Ruiz, G. (eds) *Advances in Soft Computing. MICAI 2025. Lecture Notes in Computer Science*(), vol 16222. Springer, Cham. [https://doi.org/10.1007/978-3-032-09044-7\\_16](https://doi.org/10.1007/978-3-032-09044-7_16)
29. Kazlauskas, N., Klappenbach, M., Depino, A. M., & Locatelli, F. F. (2016). Sickness behavior in honey bees. *Frontiers in Physiology*, vol 7, Article 261. Pp. 1-10
30. JACKSON, Duncan E.; RATNIEKS, Francis LW. Communication in ants. *Current biology*, 2006, vol. 16, no 15, p. R570-R574.
31. Sumpter, D.J.T., Blanchard, G.B. & Broomhead, D.S. Ants and agents: A process algebra approach to modelling ant colony behaviour. *Bulletin of Mathematical Biology*. 63, Pp. 951-980 (2001). <https://doi.org/10.1006/bulm.2001.0252>
32. Marchal, P., Buatois, A., Kraus, S., Klein, S., Gomez-Moracho, T., & Lihoreau, M. (2020). Automated monitoring of bee behaviour using connected hives: Towards a computational apidology. *Apidologie*, vol 51, no. 3, Pp. 356-368. <https://doi.org/10.1007/s13592-019-00714-8>

33. CARDINAL, Sophie; DANFORTH, Bryan N. The antiquity and evolutionary history of social behavior in bees. *PLOS one*, 2011, vol. 6, no 6, p. e21086.
34. JONES, Beryl M., et al. Individual differences in honey bee behavior enabled by plasticity in brain gene regulatory networks. *Elife*, 2020, vol. 9, p. e62850.
35. Li, H., Zhu, G., Zhang, L., Jiang, Y., Dang, Y., Hou, H., Shen, P., Zhao, X., Shah, S., Bennamoun, M., Scene Graph Generation: A comprehensive survey, *Neurocomputing*, Volume 566, 2024, 127052, ISSN 0925-2312, <https://doi.org/10.1016/j.neucom.2023.127052>. (<https://www.sciencedirect.com/science/article/pii/S092523122301175X>).
36. Amuche, C.I.; Zhang, X.; Monday, H.N.; Nneji, G.U.; Ukwuoma, C.C.; Chikwendu, O.C.; Hyeon Gu, Y.; Al-antari, M.A. Advancements, Challenges, and Future Directions in Scene-Graph-Based Image Generation: A Comprehensive Review. *Electronics* 2025, 14, 1158. <https://doi.org/10.3390/electronics14061158>
37. Wajid MS, Terashima-Marin H, Najafirad P, Wajid MA. Deep learning and knowledge graph for image/video captioning: A review of datasets, evaluation metrics, and methods. *Engineering Reports*. 2024; 6(1):e12785. doi: 10.1002/eng2.12785
38. Jinggang Yang J., Zhe Ang Y., Guo Z., Zhou K., Zhang W., and Liu Z. Panoptic Scene Graph Generation. In *European Conference on Computer Vision* (pp. 178-196). Cham: Springer Nature Switzerland. October 2022.
39. Woo S., Kim D., Cho D., and So Kweon I. 2018. LinkNet: relational embedding for scene graph. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18)*. Curran Associates Inc., Red Hook, NY, USA, 558–568. <https://arxiv.org/abs/1811.06410>
40. J. Johnson et al., Image retrieval using scene graphs, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 3668-3678, doi: 10.1109/CVPR.2015.7298990.
41. Airin A. et al., Attention-Based Scene Graph Generation: A Review, 2022 14th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), Phnom Penh, Cambodia, 2022, pp. 210-215, doi: 10.1109/SKIMA57145.2022.10029570.
42. KRISHNA, R., et al. Visual genome: Connecting language and vision using crowd sourced dense image annotations. *International journal of computer vision*, 2017, vol. 123, p. 32-73.
43. Miller G. A.. 1995. WordNet: a lexical database for English. *Commun. ACM* 38, 11 (Nov. 1995), 39–41. <https://doi.org/10.1145/219717.219748>
44. Ward, I.R.; Joyner, J.; Lickfold, C.; Guo, Y.; Bennamoun, M. A practical tutorial on graph neural networks. *ACM Comput. Surv.* 2022, 54, 1–35.
45. Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, Maosong Sun. Graph neural networks: A review of methods and applications. *AI Open*. Volume 1, 2020, Pages 57-81, ISSN 2666-6510, <https://doi.org/10.1016/j.aiopen.2021.01.001>.
46. Lingfei Wu, Peng Cui, Jian Pei, Liang Zhao. *Graph Neural Networks: Foundations, Frontiers, and Applications*. Springer Nature Singapore Pte Ltd. 2022. ISBN 978-981-16-6054-2 (eBook) <https://doi.org/10.1007/978-981-16-6054-2>
47. Lorenz, E. N. Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?. December 29, 1972 . *Hvad er matematik? B, i-bog* ISBN: 978 87 7066 494 3. Available in: [https://static.gymportalen.dk/sites/lru.dk/files/lru/132\\_kap6\\_lorenz\\_artikel\\_the\\_butterfly\\_effect.pdf](https://static.gymportalen.dk/sites/lru.dk/files/lru/132_kap6_lorenz_artikel_the_butterfly_effect.pdf). Consulted at: april, 2025.

**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.