A Review of Artificial Intelligence in the Internet of Things

Cristian González García*, Edward Rolando Núñez-Valdez, Vicente García-Díaz, B. Cristina Pelayo G-Bustelo, Juan Manuel Cueva Lovelle

MDE Research Group, Department of Computer Science, University of Oviedo, Oviedo (Spain)

Received 10 October 2017 | Accepted 3 February 2018 | Published 24 March 2018

I. Introduction

This was the question that was asked by Alan Mathison Turing and to which he responded with the proposal of The Imitation Game [1]. The Imitation Game is played by three people, ‘A’ who is a man, ‘B’ who is a woman, and ‘C’ who is the interrogator and can have any of the genders.

The interrogator is in a room other than ‘A’ and ‘B’. During the game, ‘C’ can ask them any kind of question to determine who is the man and who is the woman, for example, about the length of the hair. ‘C’ knows the participants under the labels ‘X’ and ‘Y’. When the game ends, ‘C’ must say ‘X’ is ‘A’ and ‘Y’ is ‘B’ or vice versa. In this game, the tone of voice does not help ‘C’, since ‘A’ and ‘B’ must write the answers, preferably using a machine, which are sent by a teletype or through another intermediary to ‘C’.

This problem can be thought with a variant that is based on replacing the man ‘A’ with a machine. In this case, ‘C’ could ask a mathematical operation and see that the machine solved it or ask them if they play chess. However, to avoid undue comparisons in the game, the interrogator is prevented from asking for demonstrations, since in some cases the machine is better and in others, the human being is better, as well as each having different skills. For these reasons, this game is about the machine trying to imitate the behaviour of the human being, assuming it can give answers just like a human being would do it naturally. From this theory, the well-known Turing Test emerged.

However, the imitation from a machine is very difficult to be given, as Turing says from different points of view, namely: the theological,

Please cite this article in press as:
the reluctant people to know the problem (Head-in-the-sand), the mathematician, the one of the conscience, the one of the different disabilities, from the point of view of the memories of Ada Lovelace [2], from the continuity of the nervous system, from the argument of the familiarity of the behaviour and from the extrasensory perception.

For the reasons expressed from the different points of view, Alan Turing focused on what Ada Lovelace commented when she says that a man can ‘inject’ an idea into a machine. Thus, based on this, he hypothesised that he could inject knowledge if there was a space large enough for it, or at least, so that the machine could play The Imitation Game and leaving as a problem how to program the machines to overcome the game, because according to Turing, they already had the hardware capacity.

To carry out this idea, he proposed to make a child’s brain instead of trying to make an adult brain, and thus educate it to obtain the brain of an adult. In this way, he divides the problem into two parts: the ‘child’ program and the education process. He explains that we should not expect the first ‘child’ machine to come out on the first attempt and that we should teach it to see how its learning evolves. Thus, after several attempts, they would be getting better machines (or even worse), something that Turing compared with the process of evolution and that several researchers would later develop under the name of genetic algorithms [3].

Based on this, Turing talked about different experiments he carried out and the need of machines to simulate a human being, as is the case of the incorporation of random elements to imitate the possibility of failure that human beings have. To conclude, he commented on the thinking of many people that one could start making machines compete with humans in abstract activities, such as playing chess. However, from his point of view, it was better to mount the machines on the better ‘organs’ that money could buy and teach them to understand and speak English.

II. WHAT IS ARTIFICIAL INTELLIGENCE?

Computers can only process zeros and ones. However, years ago, Artificial Intelligence (AI) was born to offer the possibility of creating programs that would allow computers to learn. This was the purpose of the first computer scientists, such as Alan Mathison Turing, John von Neumann and Norbert Wiener, among others, who tried to equip computers with programs that contained intelligence, the ability to self-copy, learn and control their environment, being, the principle of this, trying to model the human brain, imitate human learning and simulate biological evolution [4].

Research in Artificial Intelligence began after World War II, possibly being Alan Mathison Turing the first to investigate in this field in 1947 [5], and publishing an article about whether machines could think in 1950 [1]. This article is where the famous The Imitation Game comes from, from which the Turing Test emerged. Both serve to know if a machine is intelligent enough to confuse it with a human being. In that article, Turing also explained how an AI system should be created, starting with the creation of a child machine that would be learning to be as an adult. Subsequently, in the decade of the 50s, many more researchers joined this field [5].

Later, John McCarthy, Marvin L. Minsky, Nathaniel Rochester and Claude E. Shannon coined the term Artificial Intelligence in 1955 in the Dartmouth summer research project [6]-[8]. In this project a stay of two months was proposed to investigate the principles about learning or another type of intelligence and see if the computers could simulate it, or, rather, with the name they coined, to investigate in Artificial Intelligence.

But what exactly is Artificial Intelligence? According to John McCarthy [5], Artificial Intelligence is the science and engineering that tries to make machines intelligent, trying to get them to understand human language and to reach problems and goals as well as a human being. However, there are some authors who affirm that the objective of the AI is to put the human mind inside a computer, although they may speak metaphorically. One of the ways to know if a machine is intelligent is the Turing Test.

To answer the previous question, one must also understand what is meant by intelligence, for which a reference can be made to [5]. Intelligence is the computational part of the ability to achieve objectives and has various types and degrees as occurs in humans, many animals and some machines. However, it is not necessary to consider intelligence to answer ‘yes’ or ‘no’ to a question, because, although this is sometimes the goal, machines can do much more, such as, for example, solving problems.

Some of the fields where Artificial Intelligence can be applied are, for example, video games, data mining, discourse recognition, natural language understanding, Computer Vision, the creation of expert systems and data classification [5]. Thus, Artificial Intelligence can be divided into several branches among which stand out Automatic Learning, Computer Vision, Fuzzy Logic, Natural Language Processing, Heuristics and Intelligent Agents.

III. THE SIX RULES OF ARTIFICIAL INTELLIGENCE

As for robotics, there are the three rules, created by Isaac Asimov in the first literature about robots, Satya Nadella, CEO of Microsoft, offered in an interview to the online magazine Slate1 an outline of six rules that should be observed by the designers of Artificial Intelligence [9].

1. **AI must be designed to assist humanity.** It needs to respect human autonomy, using collaborative robots to perform dangerous works, such as mining, thus safeguarding human workers.
2. **AI must be transparent, being aware of how the technology works and its rules.** Or as Satya Nadella said: ‘Technology will know things about humans, but humans should know about machines’. This would allow humans to understand how technology sees and analyses the world because ethics and design must go hand in hand.
3. **AI must maximise effectiveness without destroying the dignity of the people.** It must preserve cultural commitments, strengthening diversity. A broader, deeper and more diverse commitment to the population is needed because technology should not dictate the values or the virtues of the future.
4. **AI must be designed for intelligent privacy.** Sophisticated methods of protection are needed to ensure personal information to gain trust.
5. **AI must have algorithmic responsibility so that humans can undo unintended damage.** This is why it is necessary to design AI technology for what was expected and what was not expected.
6. **AI must avoid bias by ensuring adequate and representative research so that an erroneous heuristic cannot be used to discriminate.**

However, Satya Nadella also clarified that humans ‘should’ prioritise and cultivate a series of characteristics to be able to coexist with AIs, such as:

1. **Empathy** to create relationships with others and perceive their feelings and thoughts because this is something very difficult to replicate in the machines.

---

1 http://www.slate.com/
2. **Education** that will be necessary to believe and manage innovations that are not understood today. For this, it will be necessary to increase the investment in education that allows developing the knowledge and the necessary skills to implement the new technologies and solve problems that will need much time.

3. **Creativity** because this is the most coveted human skill, that machines will continue to enrich and increase.

4. **Judgment and responsibility** to be willing to accept a diagnosis or legal decision made by a machine, but from which we expect that still a human being remains the ultimate responsible for making the final decision.

### IV. Machine Learning

**Machine Learning**, is one of the branches of Artificial Intelligence, and, maybe, one of the most important subfields of AI [10]. It is based on the development of learning techniques to allow machines to learn [11], based on the analysis of data and applying some algorithms for this processing and others for decision making from the data already analysed.

One of the goals of **machine learning** is to facilitate the process of analyzing data intelligently, examples of which are spam filters [11], **Optical Character Recognition** (OCR) techniques [12], quantification of data relationships [12], recommendation systems [13], [14], computer-aided medical diagnostics [12], search engines [15], **Computer Vision** [16], [17] or **Data Mining** [18]–[20].

#### A. Taxonomy of Techniques According to the Type of Learning

**Machine Learning** algorithms can be classified based on the type of **feedback** received. In this subsection, we present a total of six categories: supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, transduction learning and multitasking learning.

The first is **supervised learning**. These algorithms infer a function by means of labelled training data that they receive as input [12], [21], [22]. An example of them is the algorithms used to classify images by using Computer Vision techniques. These algorithms must be fed with images labelled, for example, as correct or incorrect. In this way, the algorithm is fed so that it learns to classify the images according to the previously labelled images supplied. Therefore, it can be said that the algorithm learns for itself [23]. Some of the algorithms found in this group are some of those belonging to **Artificial Neural Networks**, **Bayesian Networks**, **Support Vectors Machines** and **Decision Trees**.

Another taxonomy is **unsupervised learning** [22], [23]. The difference with the previous one is that the input data is not labelled, and it is the system itself that must infer algorithms or patterns to recognise and label the same type of data. In this group, we find some types of **Artificial Neural Networks**, the **Association Rules** and some **Grouping Algorithms**.

In contrast, in **semi-supervised learning**, a combination of the two previous algorithms is performed [24]. In this way, the system must consider both the tagged elements and the untagged elements.

Another type of learning is **reinforcement learning** [25]. This is based on the algorithm learning based on the world that surrounds it. In this case, the algorithm learns from the answers given by the outside world based on its actions, criticizing them if they were good or bad. Therefore, this is an algorithm based on the classic trial-error learning.

An algorithm similar to the supervised learning is **transduction learning**, presented in [26]. The difference is that, in this case, the function is not created explicitly, but rather the idea is to predict the categories of future examples based on the input examples. In this way, this learning tries to create groups by labelling the elements of the received inputs. An algorithm found in this category is the **Transductive Support Vector Machine** (TSVM) [27], [28].

The last taxonomy is **multitasking learning**. These algorithms use previously learned knowledge in similar problems already experienced to improve the learning of the problem to be solved [29].

#### B. Taxonomy of Algorithms Used in Machine Learning

There are different types of algorithms to perform the learning that can be grouped according to the type of output of the same. In this subsection, we will talk about **decision trees**, **association rules**, **genetic algorithms**, **Artificial Neural Networks**, **Deep Learning**, **Support Vector Machines**, clustering algorithms and **Bayesian networks**.

The first is the use of **decision trees** [19], [20], of which information on their use in Artificial Intelligence can be found since 1986 in [30]. **Decision trees** are a set of nodes arranged like a binary tree. In the intermediate nodes, there are the necessary conditions to reach the leaves, where these conditions or predicates may imply one or more characteristics of the element. Meanwhile, at the end of each branch, what is known as a leaf, there is the output that the object must return after analyzing the different possibilities obtained in the input and traversing the tree, that is, the leaves contain the decision to be made. Based on this, these algorithms map possible entries and the possibilities of such entries that the object can contain. The operation of the decision trees to classify an element consists of beginning to analyse that element by the root of the tree and, based on whether it fulfils or not the predicate of the node in which it is found, move that element to the left or right child of the current node, repeating the process until it reaches a leaf that returns the result of the tree.

**Association rules** contain those algorithms that try to discover associations between variables of the same group [19]. In this way, they allow you to discover facts that occur in a certain set of data. A possible example of these is that, if someone publishes a tweet with a photo and the hashtag ‘selfie’, then the photo was made with his mobile. A real case is the one studied by [31], where they discovered that 90% of the purchases that contained bread and butter also had milk.

Another type of algorithms are **genetic algorithms**. These are nothing more than a heuristic search process that attempts to simulate natural selection following a process of evolution of individuals through random actions. Thus, these algorithms contain methods such as mutation, crossing and selection to find the best solution to a given problem. Some examples of the use of these algorithms are in the field of automated designs of equipment and machinery, in game theory, in the learning of Fuzzy Logic rules, natural language processing and prediction, among others. The creation of genetic algorithms is due to John Holland and his students at the University of Michigan, who developed them between the 60s and 70s and presented them theoretically in [3]. In this way, his goal was not to search for algorithms to solve a specific problem, but to import into computers the phenomenon that occurs in nature by which it adapts to what happens around it [4].

Another case is the **Artificial Neural Networks** (ANN) [22], [32]. ANNs allow automatic learning by simulating the neuron system of the nervous system of animals. Within this system, the neurons collaborate with each other to produce the output, considering that each connection between neurons has a weight, which is adapted according to the experience collected throughout its use. ANNs consist of an input layer, one or more hidden layers and a final output layer, all composed of neurons. Therefore, neural networks can learn and are used for a variety of tasks such as translation of languages or learning patterns.

**Deep Learning** is a set of algorithms that are based on modelling high-level abstractions composed of multiple non-linear
transformations [33], [34]. The use of Deep Learning allows learning data representations. For example, an image can be represented in many ways, but not all of them facilitate the task of face recognition, so, what it does is try to define which is the best representation method to perform this task. Some examples of its use are in Computer Vision and in the recognition of speech and music [34]. The algorithms used can be both supervised learning algorithms and unsupervised learning algorithms [35]. The deep learning algorithms have different internal transformations between the input of the data and the output response, each transformation consisting of different weights and thresholds that can be trained [35]. However, according to [35], there is no definition that exactly says the limits between Deep Learning, Shallow Learning and Very Deep Learning, and they propose that it should be considered Very Deep Learning when it exceeds 10 levels of depth.

Support Vector Machines (SVMs) [36], also used in similar investigations such as Support Vector Networks (SVN) [37] and Support Vector Regression (SVR) [38], are a set of supervised methods that are used to solve problems of classification and regression, producing non-linear limits by using linear limits and transforming the feature space version [18], [22]. The operation of this type of algorithm is based on the selection of a small number of critical limit cases, known as support vectors, to build the discriminant function to separate existing cases as much as possible [18], [20]. For example, by using SVM you can classify a set of images from a previous training with tagged images to teach the algorithm how to classify. An example of this is the optical character recognition presented in the first article that dealt with them [36].

Clustering algorithms are those that classify similar data based on certain criteria in subgroups, also called clusters [18], [20], [39], [40]. The success of clustering is measured based on how useful this clustering is for humans [18]. There are many different algorithms to perform this task because the notion of ‘cluster’ is not defined very precisely [41]. To do this, it looks for the distance or similarity of the data according to the defined function. These types of algorithms are commonly used in data mining [18], [20], [42], as well as in other cases of statistical analysis [43]. These algorithms are unsupervised learning algorithms.

Another type of algorithm is Bayesian networks [8], [44]. Bayesian networks are probabilistic models that represent random variables and their unconditional independencies in a directed acyclic graph. That is useful, for example, to represent probabilistic relationships between diseases and symptoms. In this way, we can calculate the probability that a disease is present or not in the body.

V. COMPUTER VISION

Within Artificial Intelligence, one of the fields is Computer Vision. It is the field that allows computers to ‘learn’ to recognise an image and the characteristics of this image. The goal that Computer Vision seeks is that machines can understand the world [45]. Getting the machines to be able to recognise elements by their image and discern what they are, calculate how far they are and their place, and all this in a quick way. However, performing this task presents many challenges.

The first challenge is how to model the objects so that the computer can discern the different types of objects to differentiate what is a person or what is a car, for example. However, these can be viewed from many different positions. Here comes another problem that is to recognise or differentiate if a person is sitting, standing, behind a vehicle or crouched, or even if the car is in front or profile. In addition, there are different light stages throughout the day, not only day and night, but also cloudy or sunny. This creates another problem that is that the same object may be in the same position that the computer recognises but that it is not able to recognise it due to light changes.

Another challenge is the one that arises about how to process the image. Related tasks must be quick and efficient searches of the image should be performed, which must be processed in different ways to normalise it and to avoid some problems previously discussed.

Currently, one of the most important fields in Computer Vision is the detection of people [46]. This is due to the boom to create intelligent systems, mostly for cars, to detect pedestrians to try to avoid or at least to reduce the severity of an accident. Some even try to recognise human movement [47].

A. General Concepts

Every part of the image counts. We must work with the colour, the size of the window that we will use to analyse the image and the intensity of the image. Therefore, the same image can be differentiated in the colour of the light and the sensitivity of the camera. Any variant will make the same take a different image, although they are the same components in it. It must also be borne in mind that the material of the surface of the different elements affects the refraction of light, which also affects the images made. All this is because the cameras we use are based on the human eye because they have three types of sensors, one for each cone existing in the human eye. The cones are photoreceptors found in the retina and there is one for each wavelength: Cones-S (short), Cones-M (middle), Cones-L (long) [46]. Thus, the sensors equivalent to these cones are the red, green and blue sensors (RGB).

Sometimes we also work with other types of images such as near-infrared images (RGB-NIR), which add a non-visible channel in the shortest wavelength of the infrared spectrum. In other cases, thermal images can be used, which reproduce the correlation between temperature and infrared emission of objects. Another type of image is depth images (RGBD) that allow the devices to measure, for example, a Kinect II, the depth of the scene with respect to the sensor.

The process of Computer Vision consists of several steps to be performed [16], [46], [48]. In Figure 1, a possible diagram corresponding to these processes is shown. The first step to create a computer vision system is to perform the features extraction to build the descriptors of the objects in the images. An example of a descriptor is the colour of the pixel, which would allow defining the object as a region of connected pixels that have the same colour. However, the descriptors usually contain tens or hundreds of different characteristics to be more stable and better. This leads to performance problems and slow scanning of the images.

![Figure 1: Cycle diagram of creation of a Computer Vision module.](image-url)
One of the problems that arise is when there are changes in the intensity of the images to process since a minimum change in it will make the color different and difficult to be recognized. Therefore, the first step should be to process the images to eliminate the effects of intensity and color changes and create descriptors using a color invariant to the intensity and the effects of color. This prevents an image taken with a different intensity or color changes due to light or camera from being different for the computer vision system’s classifier.

The second step is the generation of candidates. In this step, the regions of interest in the image are defined to solve the problem of the location of the object within it. To do this, the image will be analysed in search of the previously defined objects and will be boxed with a window within the step known as window generation. In this generation, there are multiple ways to create windows and how to apply them, such as the sliding window and the sliding window pyramid (multiscale). This helps to solve one of the problems that must be solved: different size that objects can have and the depth of which they can be found.

The next step is the classification of candidates. In this phase, we must study and verify that the descriptor is effective for the different instances of the object that we want to detect and verify that it is different from the objects that we do not want to. Thus, we must also create the border that will define what our object is and what is not the object we want. The problems encountered in this phase range from possible variances in the tonalities to the translation, since it can happen that the windows generated in the phase of window generation cut objects, transfer them or cut them and detect half of an object and half of a second object as a unique object. Some of the algorithms used in this phase of candidate classification are Histogram of Oriented Gradients (HOG) [48], Local Binary Patterns (LBP) [49]–[51], HOG-LBP [52], Scale-Invariant Feature Transform (SIFT) [53] and Speeded-Up Robust Features (SURF) [54] as descriptors, and Logistic Regression [55], [56], and Support Vector Machine (SVM) [36] as classification methods. In addition, one of the learning methods used is supervised learning.

In the phase of refinement of the decision, the best scoring images must be chosen to avoid duplicates due to the overlap of the created windows and to remain with the best identification of the desired object.

In the performance evaluation phase, we must check how the classifier works and assess how accurate it is and the faults it produces, that is, whether they can be ignored or not. One way to do this is by using a confusion matrix. This matrix contains the true positives, true negatives, false positives and false negatives that were identified by the Computer Vision module. This gives us the measurement of the quality of the classifier using four parameters. The first is its accuracy, which helps us to know the distance to the perfect classification. The following is its precision, which indicates the quality of the response of the classifier. The third is the sensitivity that indicates the efficiency in the classification of all the elements that are of the kind that we are looking for. Finally, the specificity indicates the efficiency of the classifier in the classification of the elements that are not of the class for which it was trained.

However, the task of obtaining a good model to recognize a given object is very difficult. It is important to use many and very good images and try to validate it with other different images assuring that the model works correctly. In addition, there is a need to create a model to solve a specific problem, because if a generic model is created, this may have a low accuracy, due to the difficulty to recognize images. To all this, it is important to take into account the different existing problems such as the quality of the images, the different ways of training it and the time it takes to process an image.

**B. Work Related to Internet of Things, the Recognition of Faces and Movements**

Regarding the work related to the combination of the Internet of Things and Computer Vision, hardly anything was found. However, there is a lot of related work that focuses on the recognition of people, faces and heads. This work is mainly oriented to vehicles, to prevent outrages or reduce the damage of these, and health, by creating alarms to notify quickly and effectively when older people fall at home.

One of the first works about vehicles is that of [17]. Authors are investigating automatic driving systems that can detect pedestrians or other vehicles and thus avoid collisions, either by notifying the driver or taking control of the vehicle before the accident occurs. This system is based on the recognition of objects that move and on the recognition of pedestrians. This is not only useful for IloT but also in Smart Cities, to create a safer city.

In contrast, in [16], they present a framework to detect faces in an extremely fast way while maintaining a good hit rate. In addition, this article presents a new way of representing images that can be computed much faster, a new classifier based on AdaBoost [57] and a new method of combining classifiers using ‘waterfall’.

Other authors focus more on the low level, as it is the case of the improvement of the analysis of RGB photos to help the classification of images. These improvements are useful in the classification of images in general, even though they present their idea using the recognition of people [46]. Specifically, they focus on the colours that are opposite, that is, they consider the red-green and yellow-blue tuples and the luminosity.

Other research uses Computer Vision in Robotics, to allow robots to recognize objects or people to be able to perform actions based on it, such as dodging objects, collecting the desired objects among several recognised objects or recognizing a person. For example, in [58] they conducted a survey of the use of Computer Vision in cognitive robots, which are those that are designed based on imitating human cognition, in order that robots can recognize objects, but in such a way that robots do not get saturated with the information of an image, because even if it is small it contains a lot of information. On the other hand, in the ‘survey’ of [59], they deal with the state of the art of Computer Vision in unmanned aerial vehicles (UAVs), because with their use, autonomy can be extended by improving control during flight and the perception of the environment that surrounds them.

Related to health we have the work of [60]. They did a study on different systems used for the detection of falls to help the elderly. That work collects different ‘wearables’ device systems that capture environmental data and vision-based systems. According to the authors of the survey, vision-based systems must be better than the other approaches. Likewise, authors classify the field in three types: space-time, inactivity or change of form and movement and 3D position of the head. Among these is the proposal of [61], which proposes a fall detection system based on the combination of moving images with its own vector, so that it takes spatiotemporal data of the movement and its occurrence into account. In contrast, authors of [62] used a detection method based on recognizing the background pixels and using different algorithms to distinguish a fall from other normal movements. The authors of [63] presented a system by which they recognised the person and surrounded them with an oval in order to base their system of recognition of falls based on the change of shape of the oval and comparing this with the previous ovals by means of the histograms that were created. As an example of the third type of Computer Vision applied to falls, this corresponds to the 3D position of the head, which is described in this work [64], and which presents a system that combines 3D cameras with models of the head of the person obtained at times when it is inactive and thus provide in conjunction with the different
characteristics of the movements, such as the location and duration of the movement, whether it is a fall or not.

One of the works that integrates vision by computer and IoT is the one carried out in [65], where they integrated a computer vision module in an existing IoT network, called Midgar [66], [67], to allow automation for actions and objects of different intelligence [68], based on what it is recognised in a sequence of photos. For example, if it is recognised that the sequence contains a person then the system is able to send an automatic warning, thus doing the work of a guard or places

VI. FUZZY LOGIC

The Fuzzy Logic term was introduced in 1965 by Lotfi Asker Zadeh [69], Professor at the University of Berkeley, as a way to deal with the problems of common sense. Fuzzy Logic emerged to solve problems that classical logic could not solve because classical logic can only deal with binary values (zeros or ones), while some problems require dealing with more values because these problems use expressions that they are not totally true or false, as shown in Figure 2. Thus, by having more than two values, one can have more states, that decisions can be made with more information and allow them to be represented by an intermediate value that is between absolute truth and total falsehood. Normally these states are known as linguistic variables, which can represent characteristics of type ‘size’, which in this case could have three possible values or even more, such as ‘big’, ‘medium’ or ‘small’.

![Fig. 2. Comparison between a Boolean Logic system and a Fuzzy Logic system.](image)

Therefore, this approach to the problem is questions of the type: How big is this? The answer to this type of questions depends on individual cognition, therefore, not everyone responds to the same, because, not all of us have the same opinion and perception of reality. It is medium, or even large. An example of this perception could be the place where each person lives, because, if a person lives in a city with a lot of skyscrapers like New York, possibly these people think that a building of 8 plants is small, but in the case of that a second person lives in a town, is possible that the second person thinks that this building is very big.

A. General Concepts

The first thing that emerged in the field of Fuzzy Logic was the fuzzy sets in [70], presented by Zadeh in 1965. Fuzzy sets are set with ‘not very well-defined’ limits that have a scale between 0 and 1 and serve to measure the degree of belief, also known as the degree of truth [8], of a certain element in that set. Then, in Fuzzy Logic, an element can belong to a set with a certain percentage and to another set with another percentage. Therefore, these serve to represent mathematically the intrinsic imprecision of certain categories of objects and with these, to model the human representation of knowledge, and thus, improve decision systems and Artificial Intelligence. Figure 3 shows an example of the definition of two possible fuzzy sets, the left one defines the temperature and the right one defines the humidity.

![Fig. 3. Fuzzy sets defined to represent the temperature and the humidity.](image)

Later, Fuzzy Logic was defined and presented in [69]. Since then, Fuzzy Logic has been linked in many investigations. A common problem where Fuzzy Logic is applied is in energy saving. However, there are a lot of examples where Fuzzy Logic is applied. The reason for this is that all of these need more than two values to represent the states. In the case of Grant [71], he used Fuzzy Logic to propose a diabetic control. In this way, he allows managing with fuzzy and ambiguous data and thus the diabetic control makes decisions as a person would do.

To make decisions, Fuzzy Logic uses controllers called ‘adaptative controllers’ [71] or ‘expert systems’ [8]. These controllers are based on ‘If X and Y then Z’ rules to mimic human fuzzy thinking. These rules represent the knowledge that guides these controllers or systems to make the optimal decision. The definition of these rules is complex and requires the use of linguistic variables because the human representation of knowledge is fuzzy. The process of taking a linguistic variable, for example, in the previous case of the ‘size’, input, and processing it by a function to say that it is the linguistic variable with the ‘small’, value, is known as fuzzification [72]. Thus, a linguistic variable is the assignment to a numerical value a word or sentence of the natural language.

Figure 4 shows the life cycle of some data in a Fuzzy Logic system. In this figure, you can see the data entry, with two values, 18 and 85 respectively. These data are ‘fuzzified’ to transform them through membership functions [70], [73]. The difference between Boolean logic and Fuzzy Logic at this point is that, while the Boolean logic can only be 0 or 1 and belong to a set, in Fuzzy Logic this value can belong to several sets since the value can be any between 0 and 1,
being a case that the value can belong in 0.3 to one set and in 0.7 to another [70]. The **membership functions** are executed both in the ‘fuzzification’ and in the ‘defuzzification’. In this way, by means of **membership functions** it can be established to which or which fuzzy sets a value belongs. The **membership functions** can have different forms, such as triangular, trapezoidal, gamma, sigmoidal, Gaussian, or pseudo-exponential functions.

![Fig. 4. Cycle diagram through a Fuzzy Logic system.](image)

Thus, from the received values, which are 18 and 85, the corresponding **linguistic variables** are obtained, which in our example are ‘small’ and ‘large’. The next step is to apply the linguistic rules of type IF-THEN, which are found in the controller called ‘expert system’, to these variables to obtain the **fuzzy number** or **fuzzy variable**. These **linguistic rules** are responsible for obtaining the fuzzy number from the combination of linguistic variables. They are defined by hand and applying a bit of common sense [73], which implies that a test-error system may have to be made until the rules that best suit our system are found. On the contrary, if the **Fuzzy Logic** system had a training mechanism with which to learn and evolve over time, it would be said that it is an **Adaptive Fuzzy Logic System** [74].

**Fuzzy numbers** have a value between 0 and 1 and are the result of joining several values. The **fuzzy number** obtained is passed through the ‘defuzzifier’, phase in which one of the many existing methods is chosen to perform this operation in order to transform the **fuzzy number** into the normalised data that we will use to make the decision.

**B. Related Work on the Internet of Things**

Below are some examples of using **Fuzzy Logic**. Larios et al. in [75] locate a device avoiding many localisation errors and, therefore, achieve better accuracy. In this case, the authors also achieved savings in the energy consumption of different elements such as the Global Positioning System (GPS). On the other hand, Chamodrakas and Martakos in [76] proposed using **fuzzy set** representation methods to select a network and connect to it in an efficient way, with low power consumption, a good quality of service (QoS) and with a good performance.

In contrast, in the proposal of Bagchi [77], **Fuzzy Logic** was used to maintain streaming playback quality and achieve improvements in energy consumption. However, there are also proposals for the improvement of vehicle systems, such as the cases of [78]–[80]. In the first, the authors proposed improvements in the exchange of information between vehicles and servers in order to save energy. Meanwhile, in the second, the authors proposed a system to create applications through the use of voice. Meanwhile, others focus on energy saving in homes, as in [81], where they use **Fuzzy Logic** to make decisions based on outdoor temperature and humidity and the interior temperature of the house using two IoT networks and to know if it is necessary to turn on or to turn off the heating and/or air conditioning.

**VII. Natural Language Processing**

Natural Language Processing or **NLP**, sometimes called computational linguistics [8], is a branch of Artificial Intelligence that tries to solve the problem of understanding natural language using machines, or in other words, allowing users to communicate with the machines in a faster and more effective way using natural language [82], [83], as if they were communicating with other humans [84]. El **NLP** requires understanding the structure of sentences, but, in addition, it also requires knowledge of the subject matter and the context in which it is found [8]. This seeks to make applications that can analyse, understand, alter and/or generate natural language according to help with tasks ranging from writing using computers to doing more human tasks through robots. For this reason, **NLP** is a field that seeks to close the existing gap in communication between humans and machines [85].

However, **NLP** is a complicated field, since it requires a linguistic corpus, a model of the domain, knowledge of the domain in which you want to work and linguistic knowledge of the language in which you are going to work, plus, of course, the software tools and necessary hardware [86]. This implies that different disciplines must work together such as computer science for methods and creation of software, linguistics to create processes and linguistic models, mathematics to identify formal models and methods, and neuroscience to explore the mechanisms of the brain and other physical activities [83].

Natural language processing has existed since the 1940s and since then it has contributed significantly in the fields of interaction with the computer, both theoretically and practically [85]. It all started in 1946 when Warren Weaver and Andrew Donal Booth discussed the feasibility of machine translation to break enemy codes in World War II. However, it was not until 1957 when it was found that solving this problem could be much more complex than it seemed [8]. Since then, this field has improved from using dictionary-based translations to what is now known as **NLP**. An example is **NLP** in medicine, where it began to be published in the year 1960 [86].

Note that **NLP** is very important not only in the branch of computer science but for many other areas. For example, in healthcare, it is very important because there is a lot of relevant information in the text that can be processed by a machine in order to help improve patient care and advance in medicine [86]. An example of this importance is the one that occurred in 2012, when the National Library of Medicine of the United States of America sponsored a workshop on April 24 and 25 to review the current state of the art, the challenges, the obstacles and the effective use and best practices of **NLP** focused on texts in English, both in language in general and in biomedical texts [86].

**A. Application Areas**

**NLP** has many application areas [82], [83], [85], [87]. Among these, some of those that can be highlighted are in the **recognition and generation of speech**, natural language interfaces, speech management, understanding of history and text generation, machine translation, and intelligent writing assistants.

In **speech recognition**, the goal is to convert the words spoken by a person through a microphone to represent them written on a machine. Thus, **speech understanding** systems try to improve the ‘understanding’ of the machine about what the user is saying while acting and/or making decisions based on that context.

On the other hand, the **speech generation** seeks to improve the speech process of the machine by means of looking for the ideal sound representation of the written words that the machine must ‘say’.

The research in **natural language interfaces** tries to close the gap between the different existing languages in order to internationalise correctly the user interfaces due to the differences between the different languages. Another example within this category is to facilitate the
work with the interfaces through voice commands.

On the other hand, **NLP** is also used to index and classify text, perform summaries, index search engines, Data Mining, make question-answer dialogue systems, extract content, among other similar applications. In these cases, we are talking about **speech management**, **understanding of history**, and **text generation**.

A case similar to the previous one, where the system must have a minimum knowledge of what it reads, is that of **machine translation**, the oldest application of **NLP**. This type of application tries to translate a text from a given language into its equivalent in another required language [8].

**Smart Writing Assistants** is another **NLP** area. These include various types of applications such as spell-checkers, word-splitting agents, format/separation/text selection wizards, automatic thesauri, and automated document creation or maintenance or document classification environments [88], between some examples. As we can see, all these types of applications are those that try to give a smart help when are working with documents.

### B. Linguistic Knowledge Models

To perform **NLP**, different **linguistic knowledge models** can be applied [83], [85], [86], [88], [89]. These models can be classified into five different types: **symbolic or knowledge-based**, **scientists**, among which are **stochastic or probabilistic** and **connectionist**, **hybrid approaches** and **Artificial Intelligence**. Each approach has its advantages and disadvantages, which makes each one adapts better or worse depending on the type of **NLP** problem that you wish to solve.

**Symbolic or knowledge-based models** are based on the explicit representation of facts about language through schemes of representation of knowledge understood, by marking the linguistic corpus used to train it, and associated algorithms [85], [86]. These techniques are the most studied in **NLP**, however, they have many deficits in comparison with stochastic and connectionist techniques. Symbolic techniques are especially used in cases where the linguistic domain is small or very well defined [85].

**Stochastic or probabilistic models** use several probabilistic techniques to develop approximate generalised models of the linguistic phenomenon, which are based on real examples of these phenomena. These types of techniques are very effective in those domains where symbolic techniques fail [85], [86], trying to compensate for the difficulty of creating a grammar that fits the whole language and is prepared for reasoning and uncertainty [86].

**Connectionist models** also use examples of linguistic phenomena to develop generalised models. However, being made from connectionist architectures, these models are less restrictive than stochastic, which is more difficult to develop than connectionists. Being based on stochastic models, connectionist techniques are used in scenarios where symbolic techniques fail [85].

**Hybrid approaches** use the combination of different architectures and previous models in order to find the best approach. An example of a hybrid approach is that carried out by Miikkulainen in [90], where he combines symbolic techniques with connectionist techniques to process short stories and make summaries and possible answers to questions. Hybrid techniques take the strengths of symbolic, stochastic and connectionist techniques to try to minimise the effort required to construct linguistic models and maximise their flexibility, effectiveness and robustness [85], [86]. This type of approach is the most used in recent years [86].

The methods based on **Artificial Intelligence** are very varied. These can use different types of learning as shown in [88], this work shows some examples where supervised or semi-supervised learning algorithms, based on decision trees, support vector machines, artificial neural networks, Bayesian networks or genetic algorithms were used.

### C. Internal Knowledge Levels of Natural Language

**NLP** is a very extensive field that needs the analysis of different points of view to get to process a speech correctly. Thus, from the linguistic point of view, **NLP** has many aspects that must be studied [91].

- **Phonetics**: studies the sound of human speech from the point of view of its production and perception.
- **Phonology**: study the structure of sound and how it works.
- **Morphology**: studies the internal structure of words to delimit, define and classify words and the formation of new words.
- **Syntax**: studies the structure of the sentence by means of the combinatorial rules and principles of the words and the relationships between them.
- **Semantic**: studies the meaning and denotation of symbols, words, expressions and formal representations.
- **Linguistic variation**: study the different styles of expressing the same meaning and the dialects of a language.
- **The evolution of language**: it studies how the current language emerged and evolved.
- **Pragmatic**: studies the use and communication of language in a specific context.
- **Psycholinguistics**: it is a branch of psychology that studies the production and understanding of language by human beings to see how they acquire and use it.
- **Linguistic acquisition**: studies how language is acquired during childhood.
- **Neurolinguistics**: studies the mechanisms of the brain that facilitate knowledge, compression and language acquisition, both written and spoken.

Based on this, **NLP** is composed of different levels. These vary according to the authors and the time, since some consider four [82] in 1993, another six [92] in 1995, other authors eight [83] in 1995, but which were updated to nine later [85] in 1998, but in general all agree that they can be subdivided mainly into lexical, syntactic, semantic, and pragmatic [85].

Each of these levels offers a different kind of information about the same data. Some tell us the sound of the word, others tell us its meaning, others its gender and number, another level tells us that it means depending on the context, among many other data that we can obtain depending on the way of analysing the text. Thus, architectures that add more levels, get more data and different than architectures with fewer levels. The architectures with more levels add, among other characteristics, the phonetics and knowledge of the world, maintaining the levels of the previous ones.

Next, all the levels found in the literature will be explained, showing in parenthesis what kind of classification they belong to. An example of the different levels can be seen in Figure 5, which shows an example of processing the text ‘The wise dragon imitates a duck’ that needs levels 4 to 6 because its lexical, syntactic and semantic composition is studied. In case you had to recognise the sentence through a microphone when it was dictated by someone, then you would need levels 1 to 3. If, on the other hand, you wanted to recognise the situation where that sentence was written in order to keep a conversation, then you would also need levels 7 to 9. If, on the other hand, you want to have a conversation with someone, you would need all the levels.

1. **Acoustic or prosody** (8, 9): study the rhythm and intonation of language and how to form their phonemes.
2. **Phonologic** (6, 8, 9): study the relationship between words and the
sound that they make and how they work, as well as the formation of morphemes.

3. Morphologic (6, 8, 9): studies how words are constructed from their morphemes, which are the smallest unit of the language that has the lexical or grammatical meaning of its own and cannot be divided into smaller units. Thanks to this, you can know and create the gender and number of a word and thus also be able to form them. However, some authors put together the morphologic and the lexical [82].

4. Lexical (4, 8, 9): works with words, by means of lexemes and the meaning of words. In the event that the morphologic level does not exist, its constituents are treated at this level as morphemes and their inflectional forms. In addition, it works with the derivation of units of meaning.

5. Syntactic (4, 6, 8, 9): works with the combination of words that form a sentence. In this step, each word has a different category such as name, verb and adjective, which makes each word also have specific rules of combination in the given language. He also works on how to form sentences. Figure 5 shows an example of a syntactic analysis of a sentence formed by a subject (first noun phrase) and a predicate (VP) which contains the verb, and the second nominal phrase (NP). Within both nominal phrases are the analysis of each word that is classified as a determinant (det.), adjective (adj.), noun, and preposition (prep.).

6. Semantic (4, 6, 8, 9): deals with the individual meaning of words, the meaning of words according to the sentence and the meaning of the whole sentence. Thus, study the meaning of the sentence independent of the context. For this, he works on how to derive meaning from sentences.

7. Discourse (8, 9): deals with the structural roles of the sentence or collections of sentences and how to form dialogues.

8. Pragmatic (4, 6, 8, 9): works with the monitoring of the context and the interpretation of sentences within this context. This affects later sentences because it influences the meaning of the interpretation of pronouns.

9. Knowledge of the world (6, 9): this includes information about the structure of the world of language users, thus including the beliefs and goals of the other user according to maintain, for example, a conversation.

As we can see in this paper, the Artificial Intelligent is a very vast field. Besides, we have seen some examples of the use of AI with the IoT and how this fusion can create very important and interesting
applications. However, we have only touched four of the most important parts: Machine Learning, Computer Vision, Fuzzy Logic, and Natural Language Processing.

These four subfields are very related to each other; mainly, by the algorithms that are used to process the different information or the techniques that are used to teach the necessary information to create the model that different techniques need. Furthermore, these subfields are used to create smarter machines or programs to help us in our daily life, like the Internet of Things is trying to do. For instance, Machine Learning could be used to create a better module with enough intelligence to automate and take decisions, Computer Vision to automate visual scenarios, Fuzzy Logic to give intelligence and take decisions to save money, electricity, and so on, and NLP to improve the understanding between machines and the humankind.

This is why the integration of these fields, the IoT with Artificial Intelligence, could open a new world of opportunities, as the related work has shown. Even though, they have been only a few works on this topic and it is necessary a lot of research in the intersection of these two fields.

References


Cristian González García

Cristian González García is an assistant professor in the Department of Computer Science of the University of Oviedo (Spain). He is a Technical Engineer in Computer Systems, M.Sc. in Web Engineering, and a Ph.D. in Computers Science graduated from School of Computer Engineering of Oviedo in 2011, 2013, and 2017 (University of Oviedo, Spain). He has been a visitor Ph.D. candidate in the University of Manchester, United Kingdom. Besides, he has been working in different national and regional projects, in projects with private companies, and in the University of Oviedo as a professor. He has published 10 journals articles, 4 conference articles, and 2 book chapters.

Edward Rolando Núñez-Valdez

Edward Rolando Núñez-Valdez is an assistant professor in the Department of Computer Science of the University of Oviedo (Spain). He is Ph.D. from the University of Oviedo in Computer Engineering, Master in Software Engineering from the Pontifical University of Salamanca and B.S. in Computer Science from Autonomous University of Santo Domingo. He has participated in several research projects, has taught computer science at various schools and universities and has worked in software development companies and IT consulting for many years. He has published several articles in international journals and conferences. His research interests include Web Engineering, Artificial Intelligence, Recommendation Systems, Modeling Software with DSL and MDE.


Vicente García-Díaz

Vicente García-Díaz is an associate professor in the Department of Computer Science of the University of Oviedo (Spain). He has a PhD from the University of Oviedo in Computer Engineering. He has been a visiting professor in the University of Manchester (UK), Distrital University Francisco José de Caldas (Colombia) and the University of Lisbon (Portugal) and taught courses on computer science in other international institutions. He has published more than 70 research papers in international conferences and journals. He has graduated in prevention of occupational Risks and is a Certified associate in project management through the Project Management Institute. His research interests include domain specific languages, data science, user of computers for learning and teaching, machine learning and extraction of knowledge from data, extracted from social networks or internet of things-based networks.

B. Cristina Pelayo G-Bustelo

B. Cristina Pelayo G-Bustelo (Santander, 1971) Computer Science Engineer from Oviedo University, Spain (2004). Ph. D. from Oviedo University, Spain (2007). From 2004 he is Associate Professor at the Languages and Computers Systems Area in Oviedo University (Spain). His research interests include the IoT (Internet of Things), Web Engineering, Modelling Software with BPM, DSL and MDA.

Juan Manuel Cueva Lovelle

Juan Manuel Cueva Lovelle (Oviedo, 1958) Mining Engineer from Oviedo Mining Engineers Technical School in 1983 (Oviedo University, Spain). Ph. D. from Madrid Polytechnic University, Spain (1990). From 1985 he is Professor at the Languages and Computers Systems Area in Oviedo University (Spain). ACM and IEEE voting member. His research interests include the IoT (Internet of Things), Language Processors, Human-Computer Interface, Web Engineering, Modelling Software with BPM, DSL and MDA.