

Shape and Context-Based Recognition of Standard Mechanical Parts in CAD Models



Brigida Bonino ^{a,b,*}, Franca Giannini ^b, Marina Monti ^b, Roberto Raffaelli ^c

^a DIME, Università degli Studi di Genova, Italy

^b IMATI, Consiglio Nazionale delle Ricerche, Italy

^c DISMI, Università degli Studi di Modena e Reggio, Italy

ARTICLE INFO

Article history:

Received 30 January 2022

Received in revised form 15 October 2022

Accepted 23 October 2022

Keywords:

CAD model processing

3D part recognition

Standard part

Assembly analysis

ABSTRACT

The work here presented is part of a wider research project aimed at extracting and using in industrial applications high level semantic information from 3D product models that are described by means of their boundary representation (B-rep). The specific focus of the paper is the recognition among the components of the CAD model of an assembly of those belonging to some categories of standard parts largely employed in mechanical industry. The knowledge of these components is crucial to understand the structure of mechanical products as they have specific meaning and function. Standard parts follow international standard in shape and dimensions, and also typical mounting schemes exist concerning their use in the product assembly. These distinctive features have been exploited as a starting point to develop a multi-step recognition algorithm. It includes a shape-based and a context-based analysis both relying on the geometric and topological analysis of a CAD model. As already anticipated by Voelcker in his visionary ability to anticipate open challenges, the shape of an object alone is not enough to understand its function. Therefore, context assessment becomes crucial to validate the recognition given by the shape-based step. It allows to uniquely recognize components in mechanical CAD models, by confirming correct results, refusing the false positives, as well as choosing the correct one when the assignment is multiple.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

The present paper is written in memory of professor H.B. Voelcker, who together with some colleagues, such as A.A.G. Requicha, V. Shapiro and J.R. Rossignac to name some, has been a pioneer in the Computer-Aided Design system area, by defining the theoretical basis, providing an early implementation with the PADL-2 modeler and in understanding its potential and shortcomings in manufacturing industries [1–9]. Voelcker, in his visionary ability to anticipate open challenges, underlined that the shape of an object alone is not enough to understand its function. The here presented work deals with this problem providing a method for recognizing functionally important parts by analyzing product shape and structure to improve efficiency in the overall product development and life-cycle assessment.

In mechanical industries, from the 1950's, with the advent of mass manufacturing, the need to automate the development process has arisen, in order to reduce the time and the costs of

production. The first main challenge was to implement systems to computationally represent mechanical parts and assemblies, as well as schemes and algorithms to automatically define the manufacturing process [1]. In this context, solid modeling proved to be the basis for the solution. It includes theories and techniques that enable an “informationally complete” and intuitive representation of solids and support the creation, visualization, and analysis of digital models of physical objects. The principles Voelcker defined and the systems he developed addressed many of the issues in three dimensional objects representation, and continue to be the central focus of today's Computer-Aided Design (CAD) and stand at the base of Computer-Aided Engineering (CAE) operations. Nowadays, in fact, the digital representation of mechanical assemblies through CAD models is taken for granted, and it is then exploited to computationally address more complex engineering tasks. In particular, research goals such as the product-process simulation, the automation and optimization of the manufacturing process, the integration of systems and the improvement in data management have arisen and have been more compelling with the advent of the digital factory and Industry 4.0 [10].

As far as the products' information is concerned, product life cycle management (PLM) systems have been developed in the last

* Corresponding author at: DIME, Università degli Studi di Genova, Italy.

E-mail addresses: brigida.bonino@ge.imati.cnr.it (B. Bonino), franca.giannini@ge.imati.cnr.it (F. Giannini), marina.monti@ge.imati.cnr.it (M. Monti), roberto.raffaelli@unimore.it (R. Raffaelli).

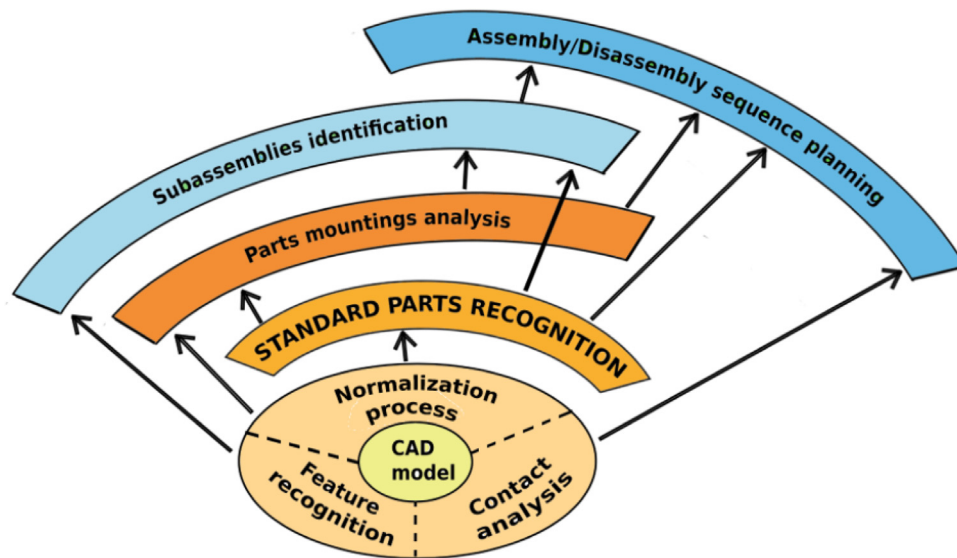


Fig. 1. Overview of the topics addressed in authors' research.

years. These are increasingly exploited by industrial companies along the various product life cycle stages. Nevertheless, not all the necessary data may be available and their inclusion is mainly demanded to the user. For example, the parts of a product may be stored in PLM systems with no indication of their meaning, or some relevant category's specifications may be missing. Even if present, textual annotations are not reliable as the same label can be used to designate different parts (e.g. bearing and bearing holder) or the notations used can depend on the designer and company conventions [11]. Furthermore, much implicit information is present in the solid definition of the product as a result of designers' reasoning and decision process. As a consequence, the definition and the development of tools for the automatic extraction from CAD models of engineering usable information is very promising and useful. In this regard, from the analysis of the product and the knowledge on the parts that compose it, it is possible to deduce much information on the operations to be carried out for simulation, evaluation, manufacturing, maintenance and recycling. For example, the presence of a screw indicates the need for drilling and threading operation during the part machining or the need of blocking tools and specific fastening tools for the product assembly [12]; it also indicates that the joined parts are firmly fixed together, so that they are movable jointly but at the same time easily disassembled allowing demanufacturing or recycling of the components. In this perspective, 3D shape retrieval and part classification are topics of interest [13,14], which have been applied in recent times in the mechanical engineering domain for parts identification [15,16].

In this context, the present paper proposes a methodology for understanding the semantics of parts within an assembly starting from the mere CAD models. In particular, the methodology foresees a rule-based approach focused on the recognition and classification of a set of specific and engineering meaningful components, i.e. the standard parts, considering both their shape characteristics and the context in which they are assembled. The identification of these parts facilitates the subsequent interpretation and processing of the CAD model and enhances its semantic value. Moreover, a further distinctive feature is that the algorithm also returns a characterization of the components through dimensional values typically used in mechanical engineering. It is worth mentioning that the approach strongly relies on the geometric modeling schemes and exploits the boundary representation (B-Rep) of a mechanical assembly and its components, underlying

the crucial role of geometry in mechanical design carried out by Voelcker in the early '80s [6].

The remainder of the paper is organized as follows. Section 2 introduces the authors' research project motivating the parts recognition method described in this paper. Section 3 briefly describes the related works. Then, Section 4 describes the major issues related to conventions adopted in product modeling in industry. An overview of the developed recognition approach and algorithms are provided in Section 5 and in Section 6. The obtained results validating the method are shown in Section 7, while a comprehensive discussion follows in Section 8, where key features and limits of the approach are pointed out. Finally, Section 9 provides some closing remarks.

2. Research goals and motivation

The work here presented is part of a wider research project aiming at extracting and using in industrial applications high level semantic information from B-rep models of mechanical products in standard format (e.g. STEP), that are thus only geometrically and topologically described. Enhancing CAD data with semantic interpretation, in fact, makes possible the development of knowledge intensive processes that leverage the engineering meaning of the parts and their relationships. In particular, in mechanical engineering, assemblies are made of many different parts including custom designed parts and standard parts. The former constitute the body of the assembly and are in general specifically created for the product under development. The latter, instead, comply generally with international norms in shape and dimensions (i.e. associated with codes UNI, ASME, DIN, ISO, etc.), hence are usually acquired from suppliers. Standard parts also have a precise role within the assembly, often related to connecting, blocking or spacing two or more custom designed parts. Moreover, considering standard parts, such as screws or nuts, fastening tools can be identified for the assembly/disassembly process. At a first approximation, dimensions and shape of the swept volume of these tools can be automatically generated from the designation of the standard components [12]. However, the distinction between these classes of parts in CAD models is neither explicit nor trivial, so the work deals precisely with this particular issue. To better contextualize the method presented here, an overview of the topics addressed so far in the authors' research is shown in Fig. 1. These are aimed at the creation of

a tool suite to algorithmically extract semantic information from CAD data and to combine it with available engineering data to ease several product development phases. Each box represents a module addressing a specific information extraction task. The arrangement of the boxes and the orientation of the connection arrows (i.e. from the inner boxes to the outer ones) indicate the data flow, i.e. that each module takes in input the results of all the connected modules that precede it. More in details, as for the knowledge extraction, given a CAD model, first a geometric analysis of the parts is carried out. It involves feature recognition and contact analysis, which are preceded by the normalization process of the boundary representations. Then, relying on both the obtained features and contact information and on part geometric information, the recognition of standard parts is performed (i.e. the subject of this work). The information on contacts between parts and the designation on designed and standard parts constitute the technological data of the assembly [12] that is then exploited for the analysis of the mounting of each pair of parts in contact. As for the applications and the leverage of knowledge, the identification of engineering meaningful subassemblies [17] and their assembly/disassembly planning is studied [18], with the final goal of defining feasible assembly/disassembly sequence plans for the whole model.

All the mentioned topics appear of particular interest in the industrial manufacturing research field and, also, the guiding idea of extracting and using parts' semantics, e.g. class of membership and dimensional attributes, is to improve existing approaches, which are still weak in this regard.

The addressed research problems are also of particular interest in industrial manufacturing due to the utmost importance of improving quality and customization degree, while maximizing the production efficiency and the ability to reuse products or parts of products in the perspective of circular economy. With this in mind, and with the widespread of the Industry 4.0 and Design for Disassembly principle, it becomes essential to intelligently automate the disassembly of products which is normally a complex and laborious operation. In this context, the ability to automatically break down, on the basis of specific characteristics, a complex assembly model into smaller subassemblies that can be treated independently of each other can significantly contribute, for example, to allow the reuse of functional components, to plan parallel production lines or to optimize the assembly sequence and human robot collaboration.

In the literature there are many methods that exploit the geometric and topological information of the model to represent the adjacency relationships among assembly's parts and geometric attributes through graphs or matrices. The subsets corresponding to the subassemblies are then identified by applying graph theory and optimization algorithms. However, geometric and topological information alone does not allow to consider all the knowledge implicitly contained in the CAD model which, if properly extracted and made available, could strongly affect the production process. It is to underline that in 3D models, the distinction between types of parts, as well as the recognition of mounting specifications is not evident. In fact, CAD models include the B-Rep representation of the components constituting the assemblies in their final configuration [19] and may lack information associated with both the composing parts and their mating conditions [20]. For instance, parts' material and features (e.g. threads and welding beads) may be not explicitly modeled [21], or may include rough part description and model inconsistencies (e.g. unexpected/missed parts interaction), due to modeling choices and numerical errors that influence the product representation [22]. The ability to analyze the topological geometric model of a mechanical assembly to automatically identify

and extract significant information on the semantics of the components becomes a fundamental step to optimize the production process under various aspects.

For example, the Design For Manufacturing and Assembly (DFMA) would benefit from the integration with such information. In fact, a weakness of this technique is that it is time consuming and in general needs human intervention to manually classify components and detect typical geometric properties, such as characterizing dimensions, weight, symmetries, thickness. The automatic identification in the CAD model of semantic data would allow to reduce human workload and improve the usability of DFMA. So far, little work has been done in this perspective (e.g. [23–25]). Similarly, for Assembly Sequence Planning (ASP) (e.g. [26,27]), using geometric information alone cannot ensure the reliability of the resulting sequences from an engineering point of view [28–31]. To overcome such limitations, recent works highlight the need to improve knowledge generated from product model (e.g. [26,27]). In fact, high level semantic information, such as the meaning and the function of parts, has proven to be extremely important both for the subassembly grouping, for assembly/disassembly tool selection and optimization, and for cost evaluation [12,32–35]. However, all these works assume this information available. In addition, the availability of a more meaningful description of CAD models and their parts would be useful in maintenance and reuse. It can speed up the product exploration, as well as support and simplify parts or subparts retrieval. At this purpose, some approaches exist in literature (e.g. [36–38]), but the automatic extraction of semantic information is rarely addressed.

Summing up, it can be argued that tools for the extraction of high level semantic from CAD models can find application in many engineering contexts, and especially the automatic identification of certain components is indeed of interest.

3. Related works

The recognition of the membership of a 3D part to a specific category of objects can be addressed either as a similarity problem or as a recognition problem. In the first case, a part is considered to belong to a specific class if it is similar to the other parts in the class; while in the second case, a part is considered belonging to a given class if it holds the characteristics peculiar to the class elements. Similarity and retrieval of mechanical components are research topics widely addressed in recent years [13,39–43]. Due to the growing size and variety of databases of 3D models of components and assemblies, as well as the fact that CAD models rarely include parts' specifications, such as category, dimensions and functionality, it is time consuming to design, reuse and manage mechanical product models from the early stages. The study and development of tools for the retrieval and cataloging of specific parts can therefore largely improve designers' activities in several manufacturing phases by allowing the reuse of previously designed parts and of the associated information (e.g. [44]).

In literature, several works exist which provide part classification methods. Ip et al. [45] tackle the problem of solid model classification providing a shape learning algorithm. A feature space is defined, where a decision tree learning is applied. Wheels, sockets and housing models are classified. Pernot et al. [46] propose the categorization of products based on shape descriptors and classify them in terms of characteristics that might affect the simplification of parts for the Finite Element Analysis (FEA). The parts distinguished are thin parts, parts with thin portions and normal. In [47] an approach for the automatic classification of mechanical CAD models according to the manufacturing process is provided. It exploits the surface curvature as shape descriptor. In [48] a

shape-based recognition approach is described that uses rules on face adjacency relations and attributes. Manda et al. [49] present a deep learning approach for the classification of engineering CAD models. The aim is to catalogue models according to functional classes. To do this, a convolutional neural network is built, which takes in input different 2D views generated from the CAD models. In [50] a shape-based retrieval method of mechanical parts is developed. It uses voxels to represent the 3D models, and divides it into several subspaces. The entropies of all subspaces are then calculated and constitute the feature vector exploited to classify the parts. Rucco et al. [51] present a supervised artificial neural network system which can classify 15 subcategories of mechanical parts. The recognition is based on different shape descriptors, such as spherical harmonics, geometric statistics, inner distances and shape distribution. A rule-based approach is adopted in [52] to recognize a set of standard parts directly from the B-rep product model. It analyzes the geometry and topology of parts and can provide good results when parts have rather distinct shapes.

It is evident that both procedural and artificial intelligence classification approaches have been implemented over the years. Both machine learning and rule-inferencing can be used for semantic enrichment. The approach should be chosen according to the nature of the problem context. While machine learning is suitable for objects with less distinct or undefined characteristics, objects with well defined and predictable features can be identified using rule-inferencing [53]. However, in general the above works can reliably distinguish models with different geometries, but can fail when models with similar shape and completely different functionalities have to be identified. In the engineering domain, in fact, parts belonging to separate classes may appear similar, as well as same shape may perform multiple functions. Thus, the membership to one class rather than the other mainly depends on the context of use of the parts in the assembly [15]. The functional and semantic classification of mechanical components considering the context is still an open issue that has been investigated in recent years. Jian et al. [54] try to overcome the limitation of traditional shape-based classification methods by exploiting non-geometric semantic information, such as tolerances, surface roughness, material and function, which are product manufacturing data, not always available in the assembly B-rep CAD model. Foucault et al. [55] propose a method that, only relying on geometric data and engineering knowledge formalization, automatically infers functional and mechanical information. In particular, conventional simplified representations and conventional interfaces between parts, typically employed by designers, are used to characterize threaded linkages and, hence, screws. The properties can be easily translated in geometric conventions, and are thus exploited as criteria for the identification of these assembly features and standard components in raw geometric CAD models. Lupinetti et al. [56] focus their attention on the identification of a single specific mechanism within assemblies, in order to enhance the mechanical function knowledge. Even if the work is targeted at the recognition of only rolling bearings, its strength is in the capability of detecting bearings independently on their design level of details. In [57] a shape-and-context-based classification is provided. The work is limited at the identification of some specific functional sets in speed reducers, but it underlines the importance of a multi-step approach. In particular, it is evident that the context of use of a part in the assembly is crucial to overcome misleading situations and not correct classification deriving from recognition methods solely based on shapes.

The paper is placed in this line of research. It provides a feature based approach focused on the recognition and classification of a set of standard parts considering both their shape characteristics and the context in which they are assembled. The identification

of these specific and engineering meaningful components is a strength of the work, since it facilitates the subsequent interpretation and processing of the CAD model and enhances its semantic value. Moreover, a further distinctive feature, that is missing in existing approaches, is that the algorithm not only assigns the assembly's parts to standard macro categories, but it distinguishes among more precise subcategories and returns a characterization through dimensions relevant in the mechanical engineering field.

4. Standard parts in assembly models

As it is stated in the previous section, in this paper, the focus is on the recognition of assembly components belonging to a set of standard parts mainly corresponding to fastener or locating components. The awareness of standard parts is indeed crucial to enhance the assembly understanding and provide information about assembly techniques.

The first idea is to automatically identify standard parts from B-rep models based on their shape, which is properly defined by the international norms. Nevertheless, this operation is not entirely reliable since shape is not sufficient to uniquely recognize a part. This is especially because the practice of idealizing the shape of some components in the assembly to lighten the CAD model is widely employed.

On the one hand, in some models, common mechanical parts can be simplified because only their overall dimensions are of interest. Among these, gears can be modeled with no teeth (Fig. 2(a)), just to represent their volume. Or, for convenience, bearings can be modeled with the balls series collapsed in a single torus (Fig. 2(b)), since, for example in producing sections from 3D models, it will always produce a circle representing the sphere as prescribed by the drafting standards. On the other hand, standard parts are usually simplified, since their appearance is well known and does not need to be fully modeled.

Also, conventional representations are largely exploited in industrial practices to geometrically represent semantic information of standard components lightening their aspect. In this context, the most significant assumptions regard threaded and deformable components' representation. It is a common practice not to explicitly model threads, rather to represent them as annotations or as an interference between a cylindrical hole and a cylindrical convex part. The hole represents the small diameter of the internal thread and the shank the major diameter of the external thread. This strategy provides components such as screws, nuts or studs simplified in shape, but maintaining their semantic features if analyzed in the context. Moreover, depending on designer choice and purpose, fasteners, especially screws, can be partially represented, substituted by a placeholder, or even totally removed from the model. In these cases features of the parts in the assembly provide hints on the presence of the fixture elements. For example, the presence of screws can be inferred by the existence of two adjacent coaxial holes (of which at least one is passing through) in two contact faces of adjacent components of the assembly. As for deformable components (e.g. retaining rings) they assume different shapes whether they are in resting or pulled position. From the modeling point of view, no specific rules exist on how to model elastic components within an assembly. Thus, they may be both represented as rested or stretched, as if they are in the extraction/insertion phase, or in an assembled setting. In this last case intersections with other components occur.

Then, another practice is to exploit 3D parts catalogs to import pre-built CAD models of standard parts, instead of modeling them from scratch. However, also in these repositories, depending of the company practice, standard parts can be represented completely defined in detail or simplified in their essential shape.

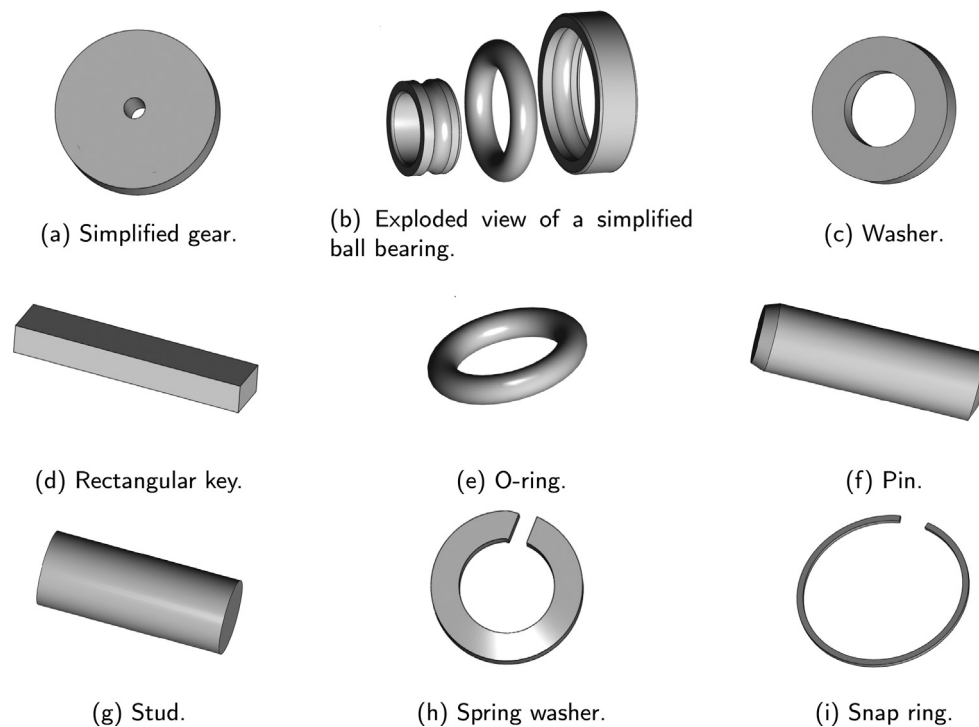


Fig. 2. Examples of CAD models of parts with simple geometric shape or simplified shape that can be misleading in the recognition.

In general, two different misleading situations can arise: a simplified component can be wrongly identified as a standard part, or else two or more distinct categories of standard parts can be confused with each other. As for the first case, just think of gears and washers that can both be modeled as simple rings (Fig. 2(a),2(c)), or else simple parallelepiped shaped components that can be mismatched with rectangular keys (Fig. 2(d)). Moreover, even multiple components can be collapsed in a single simplified component and then be mistaken for other parts, e.g. the previously mentioned balls of a bearing modeled as a torus can be mistaken for an O-ring (Fig. 2(e)). Concerning the second situation, for example, pins and studs can both be modeled as simple cylinders (Fig. 2(f),2(g)), spring washers and snap rings have both the shape of an open ring (Fig. 2(h),2(i)).

Therefore, it is evident that components with different functionality and belonging to completely different categories can appear quasi identical. This makes the recognition of standard parts very challenging.

To overcome these issues, the recognition rules have to take care of the industry widely adopted principles together with typical configurations and interfaces inside the assembly to achieve the expected connecting function. Indeed, standard parts are generally placed in specific and well recognizable seats or need to be tightened together to perform their connecting function, thus respecting precise contact relations. These considerations explain the need to define a multi-step recognition approach, which first analyze parts' shape and then their positioning within the assembly.
























5. Overview of the proposed multi-step approach

According to the above presented background, the paper provides a new approach aimed at automatically recognizing standard parts and extracting the values of their describing dimensions (e.g. diameter, length, width, etc.). That is, differently from most state of the art classification works, at the end of the process, parts are not only labeled according to their component category but, also the category-related dimensions are computed.

The method extends the work proposed in [52] by considering additional classes and enhancing the recognition through the analysis of the context in which parts are inserted. Standardization organizations have normed a large number of components covering all the production sectors. Therefore, depending on the sector, the standard parts may vary in terms of classification, defining shape and dimensional characteristics. Moreover, depending on the product class and on the materials of the parts involved, the variety of included standard elements is generally limited to some recurrent standard parts and categories. Since the research is carried out in partnerships with various Italian companies involved in the design and manufacturing of mechanical automatic machines, the presented recognition is focused on the standard parts largely employed in this type of products. The categories currently considered are: screws, nuts, O-ring, washers, circlips, keys, studs and pins. Some of these classes, i.e. O-ring and studs, refer to a single type of parts, while the others, i.e. screws, nuts, washers, circlips and pins, include a large variety of parts, which differ in features and usages. As a consequence, in the latter case, it is necessary to distinguish subcategories, in order to return more accurate results. Namely, according to practical experience, screws are divided in eight subcategories depending on the head shape, that are: *hex head screws*, *socket hex head screws*, *socket hex countersunk head cap screws*, *cross recess countersunk flat head screws*, *cross recess countersunk raised head screws*, *cross recess raised cheese head screws*, *slotted pan countersunk head cap screws*, *slotted flat countersunk head cap screws*. Nuts include two subcategories, that are *hex nuts* and *hex cap nuts*, according to whether they have a through hole or a blind hole closed on one side by a domed end. Among washers are distinguished *flat washers* and *spring washers*, depending on whether the cylindrical shape is closed or open. Circlips consist of five subcategories according to the ring ends variants and the internal shape, namely: *internal circlips*, *external circlips*, *snap rings*, *rings type G* and *rings type E*. Finally, pins are divided into two subcategories: *not holed pins* and *holed pins*.

The work presented here focuses on part categories made of a single solid body (i.e. standard parts defined as assemblies are not

Table 1
Summary table of the standard parts categories considered and the associated extracted dimensions.

Category	Subcategory	Image	Dimensions and geometric properties
Screws	Hex head		Nominal Diameter, Length, Head Height, Key Size, Center and Axis
	Socket hex head		Nominal Diameter, Length, Head Height, Key Size, Socket Depth, Center and Axis
	Socket hex countersunk head cap		Nominal Diameter, Length, Head Height, Key Size, Socket Depth, Center and Axis
	Cross recess countersunk flat head		Nominal Diameter, Length, Head Diameter, Groove Width, Cross Depth, Center and Axis
	Cross recess countersunk raised head		Nominal Diameter, Length, Head Height, Head Diameter, Groove Width, Cross Depth, Center and Axis
	Cross recess raised cheese head		Nominal Diameter, Length, Head Height, Head Diameter, Groove Width, Cross Depth, Center and Axis
	Slotted pan countersunk head cap		Nominal Diameter, Length, Head Height, Head Diameter, Slot Depth, Center and Axis
	Slotted flat countersunk head cap		Nominal Diameter, Length, Head Diameter, Slot Depth, Center and Axis
Nuts	Hex		Nominal Diameter, Head Height, Key Size, Center and Axis
	Hex cap		Nominal Diameter, Head Height, Key Size, Center and Axis
O-ring	-		Diameter, Chord, Center and Axis
Washers	Flat		Nominal Diameter, Head Height, Key Size, Center and Axis
	Spring		Thickness, Inner Diameter, Outer Diameter, Center and Axis
Circlips	Internal		Internal Diameter, External Diameter, Thickness, Center and Axis
	External		Internal Diameter, External Diameter, Thickness, Center and Axis
	Snap ring		Internal Diameter, External Diameter, Thickness, Center and Axis
	Ring type G		Internal Diameter, External Diameter, Thickness, Center and Axis
	Ring type E		Internal Diameter, External Diameter, Thickness, Center and Axis
Keys	Key type A		Height, Length and Width
	Key type B		Height, Length and Width
Studs	-		Nominal Diameter, Length, Threads Length, Center and Axis
Pins	Not holed		Length, Diameter, Center and Axis
	Holed		Length, Diameter, Hole Nominal Diameter, Center and Axis

considered), which follow international standards for shape and dimensions. Even if the list is not exhaustive, it is wide enough to cover the types typically found in mechanical assemblies having a crucial importance in the assembly process, and sufficient to prove the robustness and scalability of the approach.

The considered categories, and the associated subcategories, are summarized in Table 1, which is extending the one in [52].

The table also reports in the last column the characterizing dimensions and properties for each standard parts' subcategory that are usually considered by engineers and useful for enhancing the assembly process. The values of these parameters are extracted measuring and evaluating the geometry of the solid model of the recognized part. For instance, the key size associated with hex head screws corresponds to the diameter of the circle

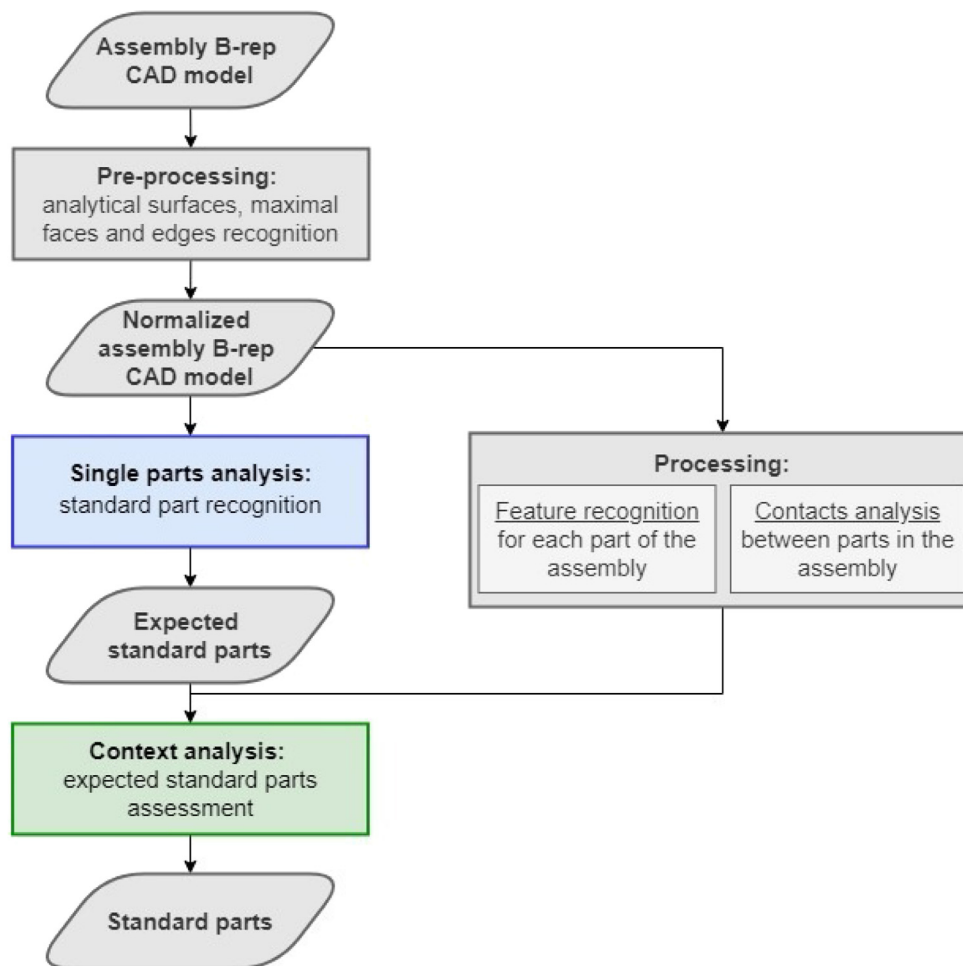


Fig. 3. General schema of the multi-step approach.

inscribed in the hexagonal plane face, the nominal diameter corresponds to the diameter of the cylindrical face of the stem, or else height, length and width of keys corresponds with the three dimensions of the model. In addition, an image is provided for each subcategory in the third column of Table 1 to better visualize the types of components. The images are representative for each class, but these can be actually modeled in alternative ways, with more or less details. However, it becomes apparent that misleading situations can occur in the recognition since some subcategories end up being quasi identical from a modeling standpoint. They are: studs and pins, spring washer and snap rings.

To overcome the issue of misclassification of parts, the method here proposed follows a multi-step approach. It consists of a single part analysis and a context analysis. Both phases are preceded respectively by a CAD model pre-processing phase and a processing phase involving all parts of the assembly (Fig. 3). The four steps constituting the proposed approach are now briefly listed, and will be addressed in detail in the following sections:

- **Pre-processing.** This first step aims to normalize the geometric (i.e. surfaces and curves contributing to the definition of faces and edges) and topological (i.e. faces and edges) entities used in the boundary representations of the assembly's parts, in order to obtain same representation for equal entities. Thus, the pre-processing phase takes in input an assembly B-rep CAD model, and returns a copy of it, where the B-rep of each part has been modified according to some general criteria (i.e. use of a single geometric formulation

for each type of surface, maximal faces and maximal edges computation).

- **Single parts analysis.** The methodology adopted in this phase is based on the work proposed in [52] that relies on a shape-based approach. In particular, this step takes as input the pre-processed CAD model and provides a part-by-part evaluation, to assess the possible membership of each component to one or more categories of standard parts. According to the engineering rules in shape and dimension followed by standard parts, geometric requirements that a part model has to meet to belong to a standard parts category are defined. Only geometric and topological information that can be derived from the component B-rep models are exploited. The output, defined as *candidate standard parts*, is a subset of the assembly's parts which have been assigned to at least a category.
- **Processing.** Before moving on to the context analysis, the CAD model is further processed. In particular, form features which are traceable to standard parts' seats are identified by feature recognition techniques, and the contacts between the parts of the assembly are computed. Starting from the pre-processed CAD model, the output of this step is the list of features associated with each part together with the list of pairs of parts in contact.
- **Context analysis.** This final step aims to consolidate and validate the classification resulting from the single parts analysis by evaluating parts with respect to the context in which they are inserted, i.e. the assembly. Analyzing

mechanical assemblies structures, in fact, it is evident that standard parts are usually arranged according to common schemes. Exploiting engineering knowledge, the mounting schemes are summarized in some rules, which are distinguished into two types. The first type of rules is associated with the relations between those standard parts generally tightened together for mounting operations; while the rules of the second type express the conditions on the positioning of standard parts in defined seats. Starting from the *candidate standard parts* and exploiting the information extracted in the processing phase, the membership of each component in the recognized category is confirmed or refused, as well as the correct category is chosen when the assignment is multiple, by verifying if the requirements are met. The context analysis returns the final list of recognized standard parts.

A distinctive feature of the approach is that both single parts analysis and context analysis are rule based. Namely, for each category shape rules and context rules have been defined that a part must respect if it belongs to that category. The combination of these two types of rules allows to uniquely identify a standard part. It is to underline that, from the point of view of industrial applications, the rule set (especially concerning sizes) can be customized in such a way to restrict the scope classification to the standard parts of interest. In the following sections the various steps are further detailed.

6. The classification algorithm

This section describes in detail the four steps of the algorithm pointing out the main concepts used and techniques adopted.

6.1. Step 0: Pre-processing

The single parts analysis relies on a shape-based approach. It only exploits geometric and topological information that can be derived from components' models. However, when dealing with CAD models in B-rep format, ambiguous situations can occur related to the non uniqueness of the representation of the parts [58]. Indeed, the same component can be described through several B-rep instances. The differences between two representations of the same model stand both in the topological decomposition in terms of faces and edges and in the adopted geometric description of the underlying surfaces and curves. In particular, a face can be divided into smaller ones that share the same surfaces and are topologically connected (e.g. a cylindrical surface can be represented either with two half cylinders or a single cylindrical face). Consequently, the same applies to edges: a set of topologically connected smaller edges laying on the same curve, rather than a single edge, can be present in the B-rep model. In addition, canonical surfaces can be described either in terms of their analytical form or in terms of NURBS patches.

In order to allow a consistent definition and application of the geometric requirements underlying the single parts analysis, it is therefore essential to restrict the space of possible representations for a certain geometry. To this aim, all the parts of the considered assembly are processed and their representation is normalized.

First, the geometry of faces is restricted to the following forms: planar, cylindrical, conical, toroidal, spherical, revolved, ruled and NURBS. For the sake of clarity, a ruled surface is obtained from a curve extruded along a linear path, while NURBS includes any level of degree, spans number and knots arrangement. Analogously, geometry of edges is limited to linear, circular, elliptical and NURBS forms. These choices derive from a practice in the most popular geometric kernels and in the neutral STEP interchange format.

The normalization of the formulation is carried out according to the following steps:

- **Face conversion:** the geometric formulation of each face, which can vary according to the originating modeler kernel and undergone file format conversions, is modified to a standardized and meaningful formulation. To make it clear, a face of revolution is checked whether it is reducible to a cylinder, a cone, a torus or a sphere respectively. Similarly, a NURBS face is verified whether it can fall respectively into a plane, a cylinder, a cone, a torus, a sphere, a surface of revolution or a ruled surface. To this purpose, algorithms provided in OpenNURBS library¹ have been adopted and the conversion is subjected to a fixed level of tolerance bounding the deviation between the original and the new form.
- **Edge conversion:** with a similar approach, edge geometries are reduced to meaningful canonical form according to the fixed tolerance.
- **Faces merging:** two faces in a body become candidates to be merged to a maximal face if they share a common tangent edge, i.e. an edge whose solid angle is 180°. Then, the underlying canonical geometry is checked to be the same, as well as the surface orientation. It applies to planar, cylindrical, conical, toroidal, spherical and revolved faces, by verifying straightforward conditions according to the specific analytical form (i.e. radii equality, points coincidence or alignment, etc.). Loops on the maximal face obtained after the merge are recomputed and appropriate classification criteria are adopted to distinguish among internal and external loops. In particular, the algorithm discriminates external loops constituted by edges of already external loops and internal loops deriving from internal loops' edges or newly originated by edges of previous external loops.
- **Edges merging:** for each loop in a face, contiguous edges (i.e. sharing the same vertex) are analyzed and merged to form maximal ones if they share the same geometric definition, i.e. lie on the same curve/line. The merge operation is accomplished by removing the connection vertex, after verifying that the vertex joins only the two edges.

The adjustment of the B-rep to maximal topological elements and to restricted geometry types allows stronger assumptions in the following recognition algorithm and limits the variability of the geometric conditions to be recognized.

6.2. Step 1: Single parts analysis

The single part analysis aims at providing a preliminary detection of standard elements in a CAD model based on their shape and the proportions between their dimensions. More specifically, the strategy adopted is to gather up engineering knowledge of mechanical components, catalogs on standards and design rules for each category and their subcategories, and then single out their most typifying aspects.

In other words, those engineering characteristics have been identified, both relative to shape and dimensions, that a component must necessarily have when it belongs to one of the categories. In addition, common rules usually followed by designers to model standard parts are taken into account.

The collected properties are then translated into appropriate geometric requirements, which are summarized in Table 2. In particular, two types of geometric requirements are considered. The first concerns the necessary presence of specific types of faces (i.e. planar, cylindrical, conical, toroidal or spherical). The latter concerns the arrangements and the relations between the faces and edges (i.e. positioning at one end of the part, symmetries, coaxiality, dimensional ranges, etc.). All assessments of the

¹ <https://www.rhino3d.com/opennurbs/>

Table 2

List of the geometric requirements for each standard parts subcategory. Let n be the number of faces of the analyzed part and let " " be the symbol that indicates the same requirement as the above cell.

Subcategory	Necessary conditions on presence of faces and their organization							
Hex nut	$9 \leq n \leq 35$	6 planar faces forming a regular polygon P_6	1 cylindrical closed face		other faces symmetric to the axis of P_6	P_6 is convex	the closed cylindrical face is concave	
Hex cap nut	"	"	"	1 spherical face	"	the spherical face is at one extrema of the BB	"	
Hex screw	$10 \leq n \leq 40$	"	"	"	"	P_6 is at one extrema of the BB	"	
Socket hex screw	"	"	"	"	"	"	P_6 is concave	
Socket hex countersunk screw	"	"	"	"	"	"	"	
Cross recess countersunk flat screw	$10 \leq n \leq 50$	4 pairs of frontal planar faces in a concave cluster C_4	"	"	other faces symmetric to the axis of the cylindrical face	C_4 is at one extrema of the BB	"	
Cross recess countersunk raised screw	"	"	"	"	"	"	"	
Cross recess raised cheese screw	"	"	"	"	"	"	"	
Slotted pan countersunk screw	"	1 pair of frontal planar faces in a concave cluster C_1	"	"	"	C_1 is at one extrema of the BB	"	
Slotted flat countersunk screw	"	"	"	"	"	"	"	
O-ring	$1 \leq n \leq 5$	≥ 1 toroidal face	other faces: cylindrical, convex toroidal, planar ring		all faces coaxial with each other	≥ 1 toroidal face convex	all the edges must be arcs of a circle	
Washer	$4 \leq n \leq 10$	2 planar faces	2 cylindrical faces		all the faces coaxial with each other	the planar faces are rings and parallel	cylindrical faces are 1 concave and 1 convex	
Internal circlip	$6 \leq N \leq 30$	≥ 2 planar faces ($\{p_i\}$ ordered by area)	≥ 2 cylindrical opened faces ($\{c_i\}$ ordered by opening angle)	2 closed cylindrical faces	p_1, p_2 are (1) anti-parallel (2) $BB(p_1, p_2) = BB(\text{part})$	c_1 is convex and c_2 is concave	c_1 and c_2 coaxial, with (1) opening angles $\geq 200^\circ$ (2) axis parallel to the normal of p_1, p_2	the 2 closed cylindrical faces concave with same radius < radii of c_1, c_2
External circlip	"	"	"	"	"	c_1 is concave and c_2 is convex	"	
Snap ring/ Spring washer	"	"	"	0 closed cylindrical faces	"	c_1 is convex and c_2 is concave or vice versa	"	
Circlip type G	"	"	≥ 6 cylindrical opened faces ($\{c_i\}$ ordered by opening angle)	"	"	c_1 is concave and c_2 is convex	"	c_3, c_4 and c_5, c_6 opened cylindrical faces, with same radius, facing each other
Circlip type E	$8 \leq N \leq 30$	"	≥ 3 cylindrical opened faces	"	"	c_1 is convex	c_1 is with (1) opening angle $\geq 200^\circ$ (2) axes parallel to the normal of p_1, p_2	2 equal opened concave cylindrical faces belonging to same surface and with same center and axis of c_1

(continued on next page)

Table 2 (continued).

Subcategory		Necessary conditions on presence of faces and their organization					
Key type A	$6 \leq N \leq 20$	≥ 4 planar faces ($\{p_i\}$ ordered by area)	≥ 2 cylindrical opened faces ($\{G_i\}$ ordered by area)	0 closed cylindrical faces	p_1 and p_2 are (1) anti-parallel (2) same area	p_3 and p_4 are (1) anti-parallel (2) same area (3) perpendicular to p_1, p_2	c_1 and c_2 convex with (1) same are and parallel axis (2) axis parallel to the normal of p_1, p_2
Key type B	"	≥ 6 planar faces ($\{p_i\}$ ordered by area)	"	"	"	"	p_5 and p_6 are (1) anti-parallel (2) same area (3) perpendicular to p_1, p_2 and p_3, p_4 $BB(p_1, \dots, p_6) = BB(\text{part})$
Stud	$3 \leq N \leq 20$	≥ 2 planar faces	≥ 1 cylindrical closed faces	"	other faces symmetric to the axis of the cylindrical face	cylindrical faces convex	only 2 planar faces are not ring and (1) antiparallel (2) at the 2 extrema of BB
Pin	"	≤ 3 planar faces	"	"	"	"	"
Holed pin	"	"	≥ 2 cylindrical closed faces	"	"	≥ 1 cylindrical face concave	"

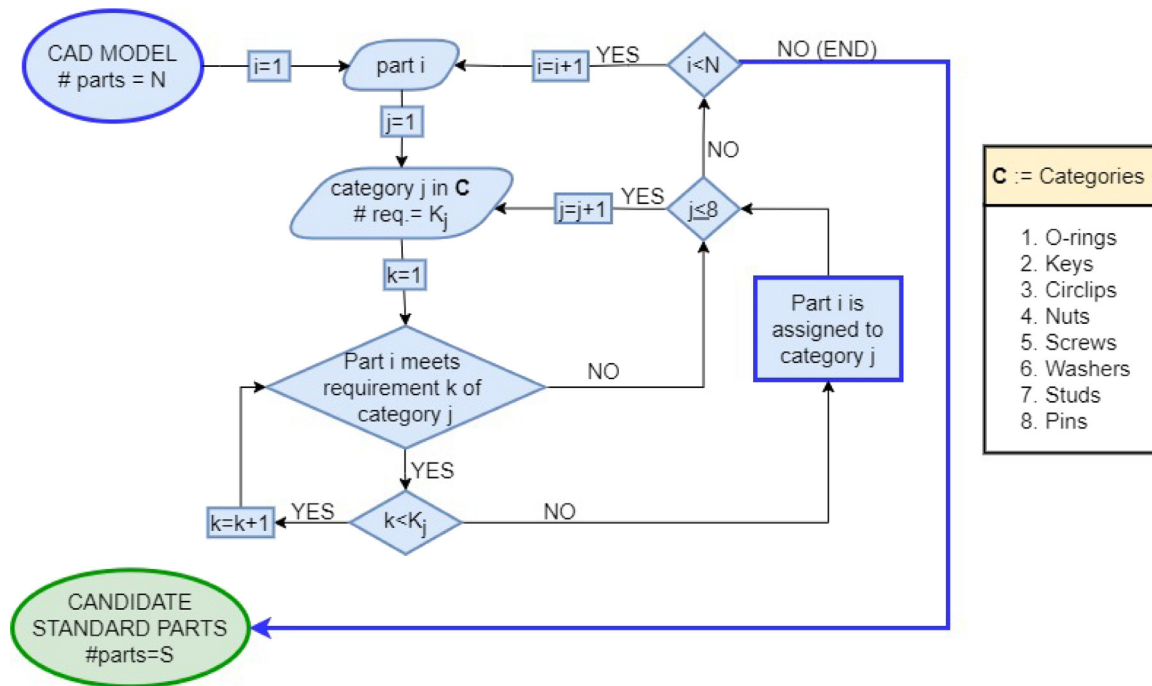


Fig. 4. Flowchart of the algorithm implemented for the single parts analysis.

relative positioning of points or axes, as well as of parallelism or perpendicularity between faces are made within a threshold, using a linear absolute tolerance equal to 0.001 millimeters and an angular absolute tolerance of 0.008 radians.

The rules are ordered in each row of the table hierarchically based on the significance of the condition. Namely, reading the table from left to right, the conditions that are more restrictive are reported first, followed by those that are less discriminatory (i.e. those related to the distinction between subcategories). This order corresponds to the one adopted in the eighth (one for each macro class) procedures for the standard parts recognition.

In the table, BB indicates the bounding box of the part computed aligned to an intrinsic frame of the part containing all its faces, e.g. in the case of axisymmetric parts the z axis corresponds to the axis of symmetry. A face, respectively a set of faces, is considered at an extreme of the BB when it, respectively a face of the set, is the face of the part having the barycentre closest to a BB's face. The $\{p_i\}$ and $\{c_i\}$ represent the sets of planar and cylindrical faces sorted in descending order with respect to the property indicated in the cells.

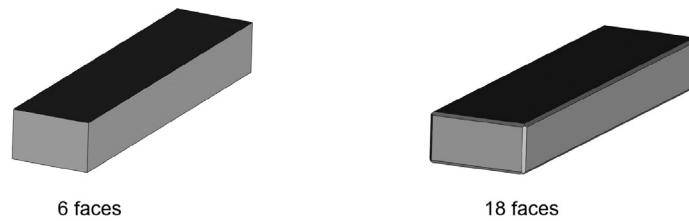
For instance, the elements p_j and p_k (respectively c_j and c_k) indicate the elements of the set with the j th and k th largest value of the considered property. For sake of space the conditions on the characterizing dimensional values are omitted.

Following the table structure, the single part analysis is accomplished as an iterative process. As shown in Fig. 4, the N parts of the assembly are investigated checking their possible membership to each of the eight considered standard parts categories collected in the set C. Namely, the identification of the i th part membership to the j th category, with $j = 1, \dots, 8$, is carried out by verifying the fulfillment of the category requirements K_j . The eight implemented recognition functions, are independent from one another, and follow a schema that, proceeding by steps, analyses the geometric requirements the part has to satisfy to belong to one of the subcategories of the given category. Once the part characteristics do not satisfy a requirement k , the function returns false, without evaluating the succeeding features, and the membership to the next category, i.e. category $j + 1$, is then evaluated. It is evident from Table 2 that the functions

can evaluate in parallel the membership to all the subcategories of the same category till the requirements are the same, but when different rules are found they split the evaluation according to subcategories adopting switch statements. To minimize the number of checks as much as possible, each function evaluates the requirements as they are reported in Table 2, ensuring that the part meets the fundamental and most representative features of the category, and then evaluates the properties associated with more precise specifications allowing the distinction between subcategories.

In general, first, the number n of faces is counted. Even after the pre-processing phase, the number of faces can vary depending on the level of detail with which the parts are modeled (Fig. 5). A minimum number of faces is mandatory, with reference to the most simplified model, and a maximum number is chosen considering all the possible chamfers, fillets and finishes. Thus, if the number of faces is out of the established range, the algorithm switches to the evaluation of the next category. If, instead, the number of faces is in the established range, the algorithm proceeds to verify the type of surfaces of the faces and to evaluate their relative positions (e.g. parallel, perpendicular), the symmetry of the part and the existence of specific faces sequences. If the requirements' combination specified for the considered category is satisfied, the component is supposed to belong to the corresponding class. To confirm this assumption, the dimensions and their ratio are checked.

In order to ensure the recognition of as many standard parts as possible, in particular also of the simplified ones, for the specification of the rules, standard parts modeled in different ways have been taken into consideration and from these the set of common rules has been derived. This choice, however, involves the fact that some categories could have similar geometric requirements, and thus can be confused. In this regard, unlike [52], a part can be associated with multiple categories, and the further context analysis is conceived precisely to confirm, refute or discriminate the obtained recognition.



6 faces

18 faces

(a) Keys of type B modeled with respectively 6 and 18 faces.



18 faces

41 faces

(b) Cross recess countersunk raised head screws modeled with respectively 18 and 41 faces.

Fig. 5. Examples of standard parts belonging to same subcategory but with different number of faces.

6.3. Step 2: Assembly model processing

The context analysis largely exploits the knowledge of some specific geometrical features traceable as seats of standard parts and the contacts between the parts of the assembly. To gather the needed information and be able to reuse it without having to compute it every time, a processing step on the CAD model of the assembly is implemented. It consists of a combination of features recognition and a contacts computation processes.

As for the features recognition, for each part, the presence of holes, pockets, slots and circular grooves is detected and stored. To this aim, procedures detecting and analyzing hollow/not hollow faces clusters are employed. More in details, holes are detected in the presence of axis-aligned hollow faces of cylindrical, conical and toroidal type, as well as disk- and ring-shaped planar

faces, and they can be both blind on one side or pass-through (Fig. 6(a)). For pockets, concave clusters with at least four sided faces and a bottom face adjacent to them are considered. Among the sided faces, there must be a pair of opposite parallel planar faces with same rectangular shape. The other two can be either opposite parallel planar faces, if it is a rectangular pocket (Fig. 6(b)), and opposite cylindrical faces, if it is a circular-end pocket (Fig. 6(c)). If the cluster includes more than four faces, conditions are checked to ensure that the other faces can be fillets. Moreover, through circular-end pockets are pockets with the bottom face missing (Fig. 6(d)). Slots are then identified as concave clusters with two or three sided faces connected by a bottom face, possibly through fillets. Slot are through when they are two sided and the faces are a pair of opposite parallel planar faces (Fig. 6(e)). Slots are, instead, blind if there is a third sided face connected

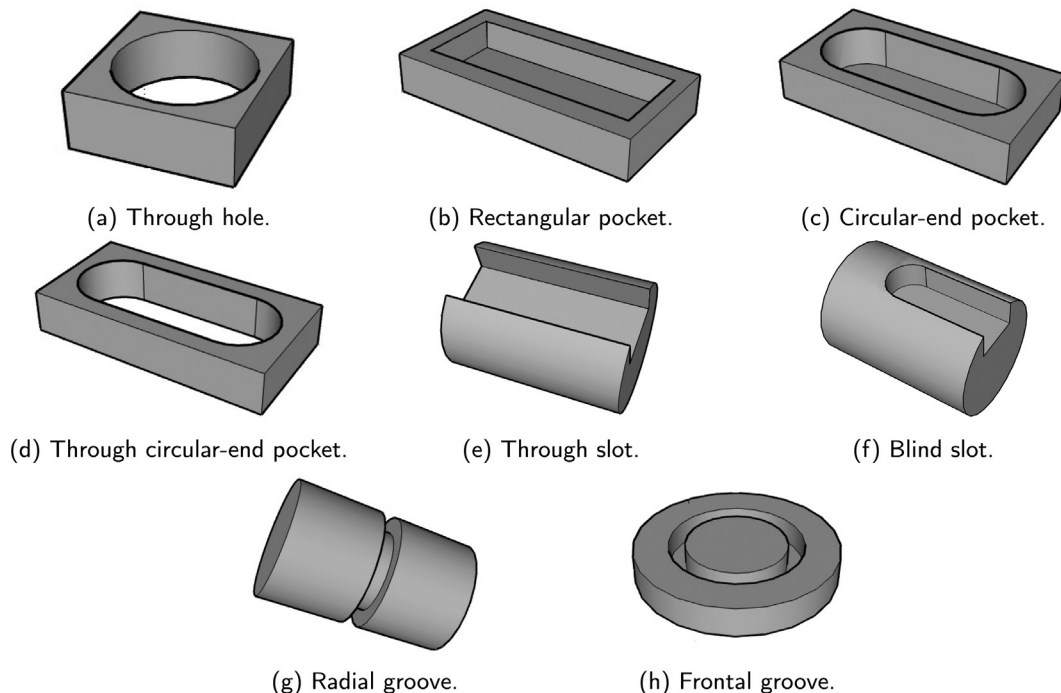


Fig. 6. Features.

Table 3

Evaluation criteria used in the context analysis: possible contacts between standard parts (in gray) and features associated with seats for standard parts (in blue).

	Standard parts in contact								Feature seats				
	O-ring	Key	Circlip	Nut	Screw	Washer	Stud	Pin	Hole	Groove	Slot	Through pocket	Pocket
O-ring										✓			
Key											✓		✓
Circlip										✓			
Nut				✓	✓	✓	✓						
Screw				✓		✓			✓			✓	
Washer				✓	✓	✓	✓						
Stud				✓		✓			✓				
Pin									✓				

with the other two (Fig. 6(f)). Finally, grooves are circular cutouts. In particular, they can be radial, when they include two equal planar rings connected by a cylindrical face (Fig. 6(g)), or frontal, when they include two equal coaxial cylindrical faces connected by a planar ring (Fig. 6(h)).

As for contacts detection, then, the surface collision between the faces of the components of the assembly is evaluated. Namely, if two parts have a pair of faces lying on the same canonical surface, partially or totally overlapped in the 2D space defined by the common surface, they are considered in contact. Contacts can be planar, cylindrical or conical according to the geometric type of the surfaces. In general, contacts are computed using an absolute tolerance threshold of 0.1. But, for cylindrical coupling the contacts detection is extended by accepting the coaxial cylindrical face pairs so that their diameters ratio is in a neighborhood of 1, precisely the maximum accepted value for the diameters ratio is 1.25, as to deal with the cases of threads interferences.

Since multiple contacts can clearly exist between each pair of parts, one for every couple of colliding faces, a list containing all the identified contacts is stored.

6.4. Step 3: Context analysis

Since the classification based solely on shape can return misleading results, due to the shape similarity among the categories' elements, the context analysis is crucial to validate the candidate standard parts recognition.

The main idea is to leverage typical engineering arrangements of standard parts in real mechanical assemblies to infer admissible characteristics or mandatory requirements that standard parts models can or have to meet. These are exploited as decision criteria to assess whether a part associated with a category from the single part analysis actually belongs to the category. As already mentioned, the extracted rules relative to the mounting schemes are of two types and target both the relations between standard parts and the positioning of standard parts in defined seats. In particular:

Relations between standard parts. In mechanical assemblies, some categories of standard parts need to be tightened together to perform their function of fixing two or more components. Consequently, recurring groups of standard parts can be identified. In

particular, these are distinctive of threaded fasteners (i.e. screws, nuts, and studs), but also not threaded components have to be taken into account (i.e. washers). Screws are threaded fasteners which have to be inserted into aligned holes of two or more mating parts to fix them with each other. Screws may be used alone, or in conjunction with nuts, to avoid axial movements and secure the connection. More than one nut can be tightened against a screw, thus two nuts may be in contact. Also washers can be mounted on screws, in general, with the function of spacer, or to distribute the load of the fastener. Studs, then, are a category of fasteners threaded on both ends. Similar to screws, studs have to be used in combination with nuts and washers to fit at least two components together.

The accepted contacts between standard parts are shown in the first block of Table 3 (i.e. checked elements in the gray columns).

Positioning of standard parts. In mechanics, it is known that some standard parts have to be inserted in defined seats to fulfill their function. Consequently, the assessment of the correct positioning of a candidate standard part can be a reliable indicator of the validity of the recognition. According to the categories we address, the corresponding possible seats can be described as hollow features carved into assembly's components (obtained by turning, drilling or milling operations). Namely, the features considered are holes, pockets, slots, and circular grooves.

Indeed, screws, as well as studs, have to fit inside at least two holes to fasten components together. More specifically, it can be affirmed that screws and studs can be inserted both in threaded holes, not threaded holes and through pockets (Fig. 7(a)). Moreover, at most one hole can be blind, while the others must be pass-through. Pins, on the contrary, are unthreaded fasteners and thus they are designed to be inserted only through preformed holes. Keys, then, are always installed on a keyseat and a keyway (Fig. 7(b)). These are referred as circular-end pockets, rectangular pockets or slots (through or blind). Finally circular grooves are the seats associated with O-ring (Fig. 7(c)) and circlips (Fig. 7(d)).

The relations between standard parts and features are summarized in the second block of Table 3 (i.e. that corresponding with the blue columns).

Therefore, for each candidate standard part in the assembly, the parts in contact with it are examined according to the decision criteria cited above. When the parts in contact are classified

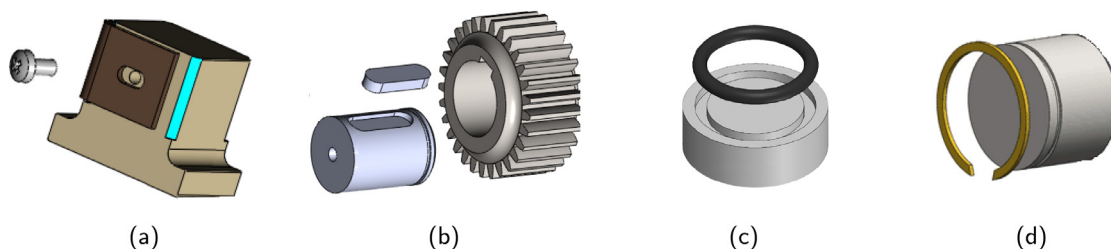


Fig. 7. Examples of standard parts inserted in the associated seats: (a) Screw in through pocket and hole. (b) Key in keyway and keyseat. (c) O-ring in frontal groove. (d) Circlip in radial groove.

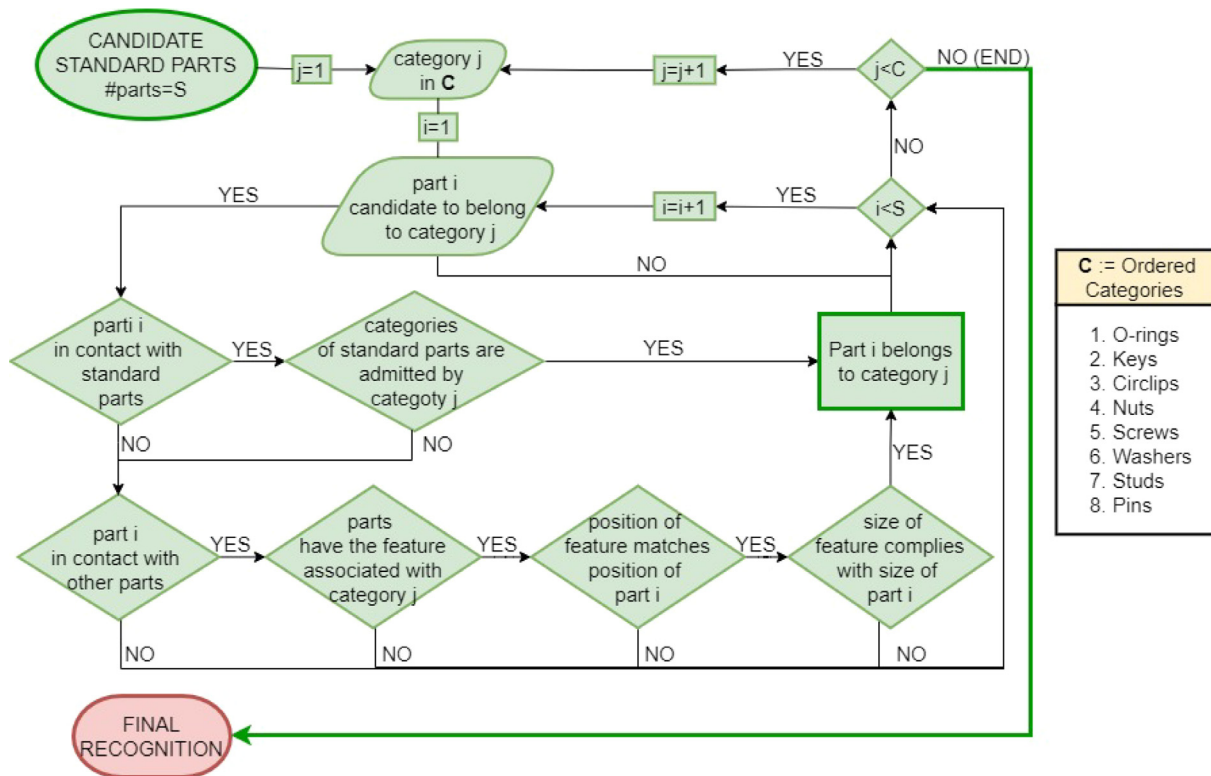


Fig. 8. Flowchart of the algorithm implemented for the context analysis.

as standard parts their category is considered, while for not classified parts the presence of specific features is verified. In order to reduce the verification steps, the identified parts are analyzed in a suitable order according to the belonging candidate category, as described in the following section.

6.4.1. Sorting of categories

The contact between standard parts is the first criterion adopted to assess the correctness of the shape-based part category recognition. However, since the classification of standard parts is the result of the shape-based classification process, its usage must be carefully pondered.

To minimize the use of misleading information deriving from wrong classification, the order in which the candidate standard parts are evaluated is crucial. In details, as shown in Fig. 8, the context analysis algorithm cycles over the categories contained in the ordered set C, which have been sorted taking into account different factors. These factors include the possible contact with other standard parts and the probability of being confused with other categories. The categories having no relation with other standard parts and a shape that can hardly be mistaken with other categories are assessed first. Taking into account Table 3, these are O-rings, keys and circlips. In this way, their validation reduces the number of uncertain data, allowing for more reliable information in the subsequent evaluations. Then, the algorithm processes the parts identified as belonging to categories relying on the contact with other standard parts for their recognition, but that are also characterized by distinctive shapes, namely nuts and screws. Washers analysis follows, since their recognition mainly depends on the relation with the just confirmed screws and nuts. Finally studs and pins are taken into account. These are two categories featured with very generic and similar shapes, both inserted in holes. In this case, the relation with other standard parts, i.e. nuts and washers, rather than the positioning in defined seats, allows to discriminate between them, therefore as reliable as possible data is needed.

6.4.2. Candidate standard parts validation

Once the contacts between parts and the features associated with the standard parts' seats has been recognized, the candidate standard parts are processed.

Fig. 8 illustrates the validation process that proceeds analyzing parts by category. Thus, once the category j is fixed, given a part i , with $i = 1, \dots, S$ where S is the number of candidate standard parts returned by the single part analysis, if part i is a candidate to belong to category j , all the relations with its adjacent parts are analyzed. It must be specified that in the algorithm a general concept of contact between the parts is used. Thus, two parts are considered in contact, or adjacent, not only when they share the contact detected as described in 6.3, but also when intersect. In this way, it is possible to overcome problems arising from modeling or numerical errors, i.e. bad positioning or sizing, or modeling choices, as, for example when dealing with deformable components (Fig. 9).

If the current category j allows contacts with other standard parts, adjacent standard parts are taken into account. If these belong to the admissible categories for the component category being validated, the membership of part i to category j is confirmed and the algorithm proceeds to next parts evaluation. If instead the adjacent standard parts belong to not admitted classes, or no adjacent standard parts are found, then the part positioning is verified.

The non-standard parts adjacent to the analyzed candidate standard part are now considered. For each of them, it is checked whether there is at least a feature associated with the current category j . However, the identification of a right feature on an adjacent part is not sufficient to confirm the class of the candidate standard part. The feature, in fact, could be the seat of another standard part, just as it could be a feature not associated with any standard part i . Further verification is needed regarding the matching of the position and the dimensions of the feature with those of the candidate standard part. First, the feature's faces

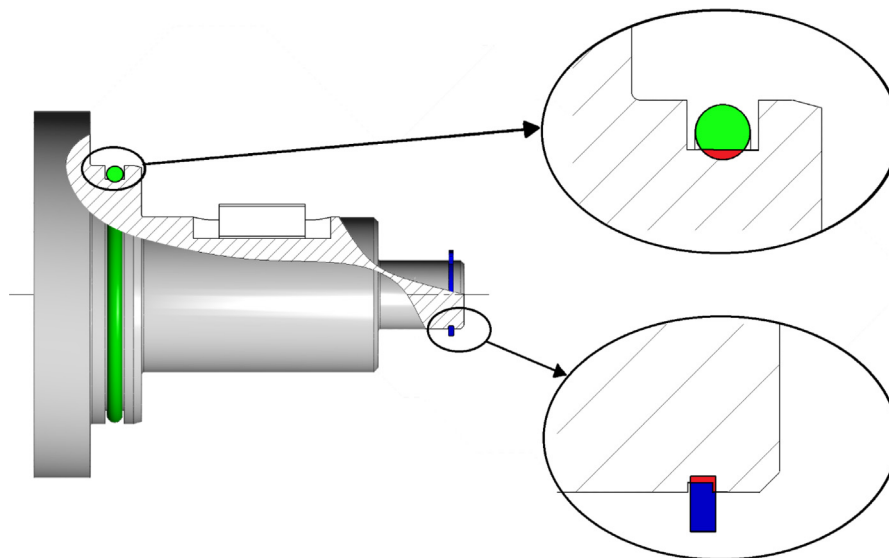


Fig. 9. Examples of deformable standard parts, i.e. O-ring (green) and circlip (blue) that intersect the faces associated with the respective seats, i.e. grooves. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

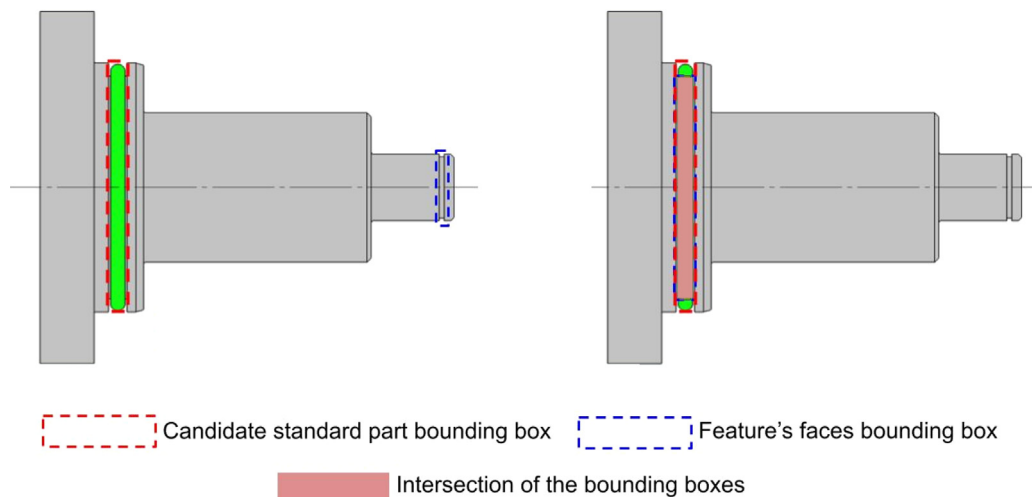


Fig. 10. Example of the matching of the position of a candidate O-ring in respect to two grooves through bounding boxes intersection (2D view).

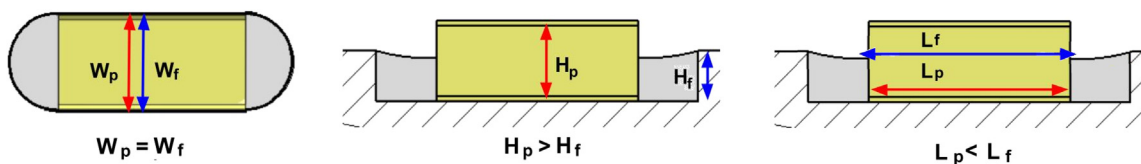


Fig. 11. Example of the matching of the relevant dimensions of a candidate key with those of the intersecting circular-end pocket.

and the candidate standard part bounding boxes must intersect (Fig. 10). If this condition is not met it means that the part is not included in the found feature and the classification is rejected. If instead the bounding boxes intersect, then the matching of the dimensions follows.

Since the candidate standard part must fit in the seat, some dimensions should coincide. The match is assessed within a threshold, in order to handle numerical and/or modeling errors, as well as the different modeling of deformable parts. The threshold is set on the ratio of the dimensions (i.e. dimension of the feature over that of the part or vice versa), which must be, in general, in the range [0.9, 1], but the range is extended to [0.8, 1] for

parts associated with grooves due to the wide variation in their representation.

The relevant dimensions (e.g. width, length, height, diameter) of a part that must respect the ranges dictated by the feature size depend on the category analyzed. The example of a key is reported in Fig. 11. It shows that it must be inserted between a keyway and a keyseat, which are identified as circular-end pockets, rectangular pockets or slots. The width W_p , the height H_p and the length L_p of the candidate key have to be compared with those of the adjacent feature, respectively W_f , H_f , and L_f . In particular the width must coincide with the feature's. The key's height must be greater than the feature's, to ensure the key be

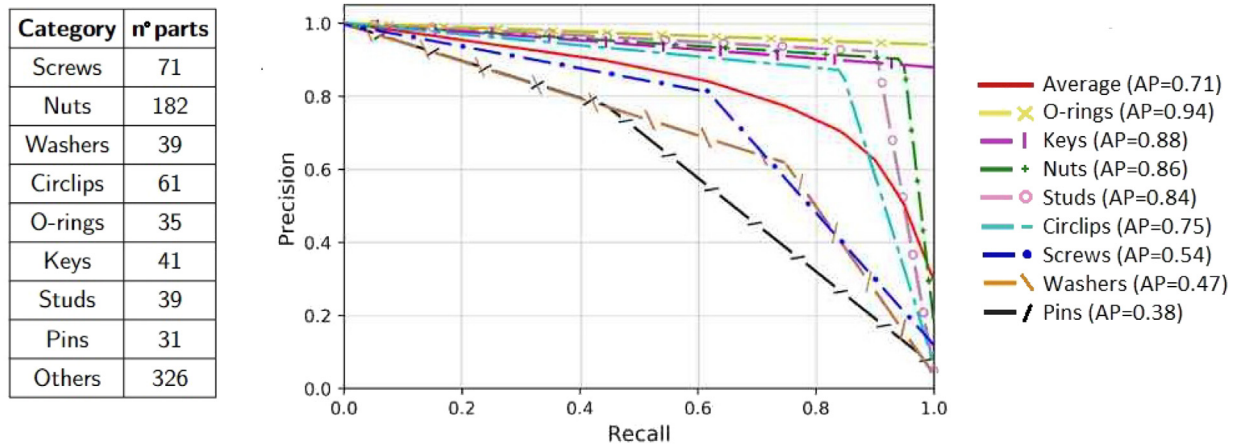


Fig. 12. Dataset organization and plot of the precision–recall curves for each standard parts category.

linked also to another part. Finally, the length of the key must at most correspond to the feature's length, that in this particular case of a circular-end pocket corresponds with the length of the planar sides, otherwise it means that the key cannot be contained in the feature.

In conclusion, as shown in Fig. 8, if a candidate standard part i meets the requirements of contact with other standard parts, or the condition of being positioned in a specific seat with compliant dimensions, it can be established that the candidate standard part was correctly classified by the single parts analysis. Its category j is confirmed, and if other categories have been assigned to the part, they are discarded. The checked adjacency relations for the considered categories are summarized in Table 3.

7. Validation

This section summarizes the results obtained to validate the presented recognition algorithm showing examples of its application on real industrial CAD models. Since the research is part of a project carried out in partnership with the Italian company Hyperlean,² the proposed algorithm is implemented in a prototype module of their industrial software LeanCOST, developed using C++, C # and VB.NET. languages.

The algorithm is tested in two stages. First the single parts analysis is assessed on a large dataset of single components, then the complete approach, given by the combination of single part analysis and context analysis, is evaluated on a set of industrial CAD models.

7.1. Single part classification validation

The single part analysis is applied on an ad hoc collected dataset of 825 mechanical parts downloaded from different online catalogs. The dataset is publicly available at <http://standardPartRecognitionDataset.ge.imati.cnr.it>. It includes CAD models in STEP format representing both parts of the considered categories and not standard parts. Parts are organized in 22 directories, i.e. one for each of the considered (sub-)categories plus the “miscellanea” one in order to provide the ground truth to evaluate the results quality according to the precision–recall metric. For sake of clarity, in the following discussion and pictures, sub-categories results are grouped into their macro categories.

In Fig. 12 the precision–recall curves for each category are plotted, and the average precision (AP) values are reported. The

average curve is provided to evaluate the overall quality of the recognition algorithm. In general, a large area under the curve represents both high recall and high precision, where high precision relates to a low false positive rate, and high recall relates to a low false negative rate.

From the graphic it is evident that some categories are well recognized, others less. The curves associated with O-rings and keys are linear and delimit a large area under themselves. This is because those classes are difficult to confuse with other standard parts classes, and their geometric requirements are few and simple. The curves of nuts, studs, and circlips are not linear, but the area under them is quite large, that is their recognition is however well performing. The remaining categories' curves instead decrease faster and indicate a worse recognition quality. This is justified by the fact that these categories can both be mistaken with other standard and non standard parts. For example, misleading interpretation of shafts as pins and/or studs can occur. Studs can be mistaken with pin, thus increasing false positives associated with pins, while the vice versa happens less frequently. Spacers and cylindrical parts can be wrongly recognized as washers, as well as spring washers are assigned both to spring washers and snap rings, and vice versa. Hex screws and cap nuts can be confused in some cases, or else they cannot be recognized due to modeling errors such as dimensions out of standards or cluster of faces included in the heads badly modeled. The average precision–recall curve (red) summarizes these results and indicates that the single part analysis approach provides good results, but not optimal. Thus confirming that shape criterion alone is not robust enough to uniquely identify standard parts especially because false positives are returned.

Finally, it is to underline that this validation results and the different values of average precision for each category in general reflect, and therefore justify, the considerations made to sort the standard parts categories before computing the context analysis (see Section 6.4.1).

7.2. Context analysis validation

While for single parts, the creation of the ground truth for the evaluation of the method can largely exploit the classification provided by professional repositories for engineers, in the case of assemblies so far no public dataset exists suitable for our purposes. This is because it is generally hard to find online repositories providing realistic mechanical models. In most cases, assemblies are made of few parts, can be inaccurate or too simplified, and thus they have shapes inconsistent with industrial practice. Moreover, the main difficulty is the limited availability

² <https://hyperlean.eu>

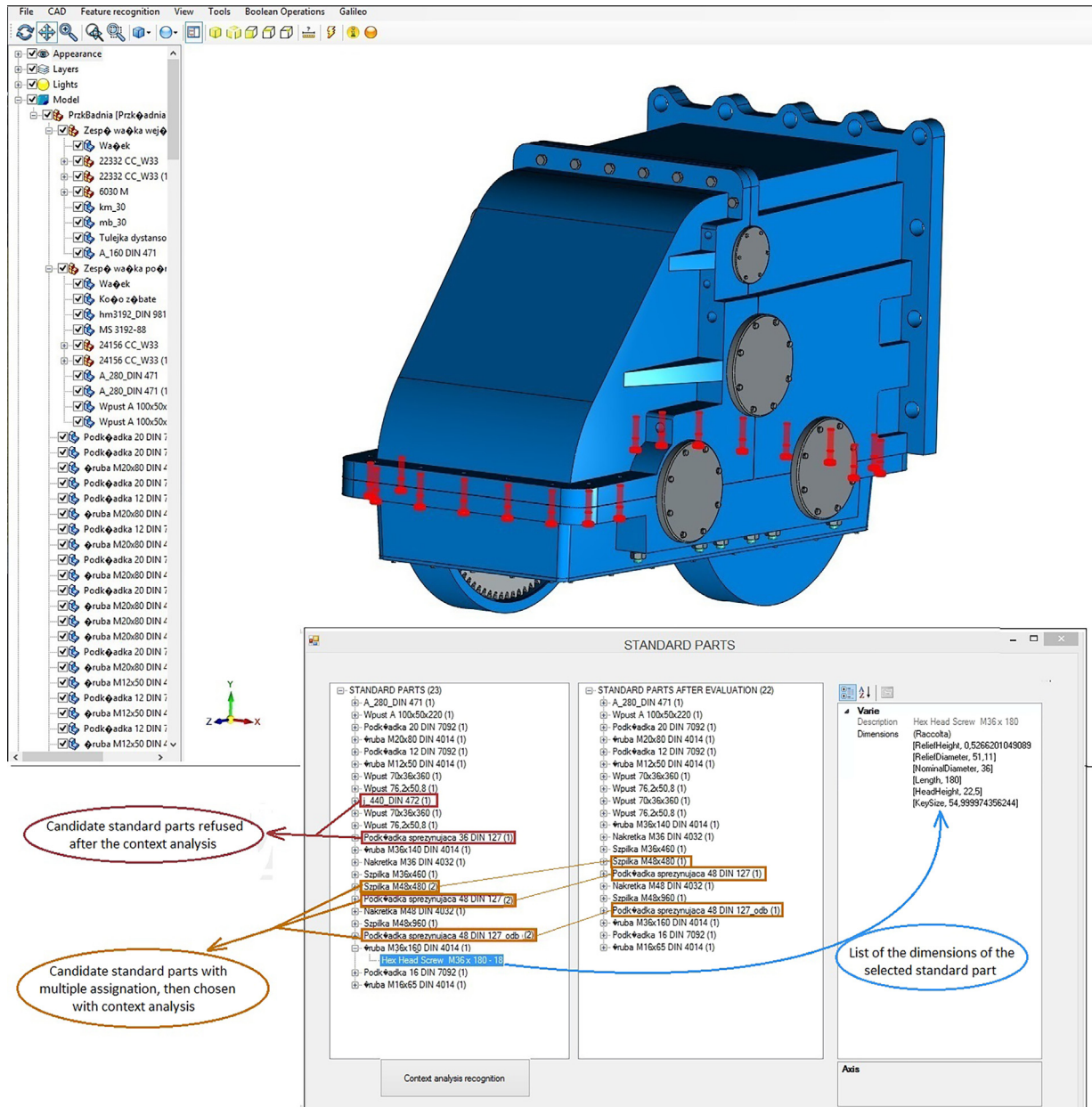


Fig. 13. Visualization of the multi-step recognition results in the developed tool.

of opportunely labeled models indicating the included standard parts. This lack prevented the creation of a meaningful, in terms of number of elements, ground truth. Consequently, to validate our work, we collected a set of industrial CAD models that contain the various categories of standard parts and can cover all the various cases addressed. Then, the results obtained by the presented method were assessed with the help of engineers.

In Fig. 13 an example of how the developed tool displays the classification results is presented. In the *STANDARD PARTS* form, the candidate standard parts and the final recognized standard parts are listed in the first and in the second column respectively. In the first list the non confirmed candidate parts are also highlighted. In both columns, the identified components are grouped according to subcategory and dimensions, i.e. all the repetitions of the same standard part are grouped together. By clicking on one item the associated components are highlighted in the CAD model to provide an overview of their arrangement within the assembly.

Finally the extracted engineering dimensions of the elements of the set are provided on the third column of the same form and an example of the dimensions extracted for a set of keys is further shown in Fig. 14.

In Table 4, some of the industrial assemblies used to validate the approach are collected. The results obtained with single part and context analysis are organized in the table as follows. The fourth and the fifth columns show respectively the number of candidate standard parts recognized after the single part analysis and the number of standard parts recognized at the end of the multi-step approach. The two columns report both the total and the per-category number of recognized standard parts recognized. That is, a set contains all the parts equally classified, both in terms of subcategory and dimensions. The last column, then, provides the number of a candidate standard parts/sets that were rejected/reassigned emphasizing the importance of the context analysis. The green and the red colors indicate the ones correctly and wrongly modified.

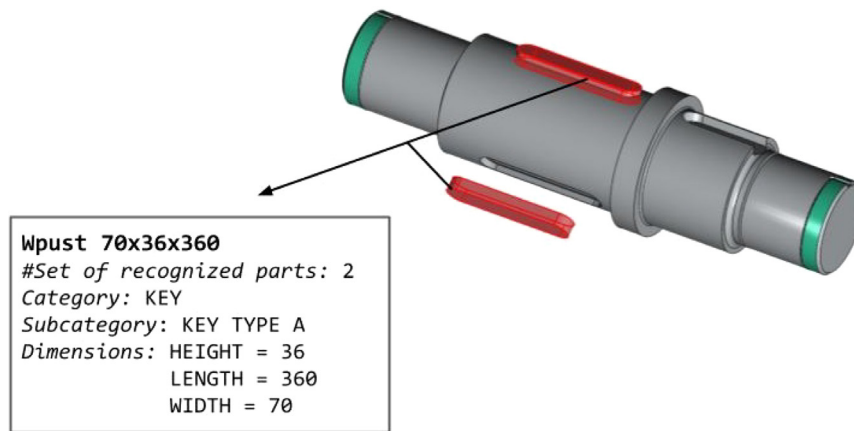


Fig. 14. Example of dimensions extracted for a set of two keys.

Results show that the provided multi-step approach is overall robust, and can recognize and correctly classify most of the standard parts included in mechanical assemblies.

Comparing the list of candidate standard parts returned by the single parts analysis and the final list of standard parts, it is evident that the context analysis allows to overcome the issues arising from the evaluation of the shape only. In particular, in the first step, classes such as studs, keys, O-rings and washers are over-recognized. That is, several false positive are returned, due to the existence in mechanical assemblies of simple components that can be mistaken for standard parts if considered out of context. Looking at Table 4, these categories of standard parts are those that actually appear most among the refused items.

For example, the case of the radiator in model N.1 is significant. From the table it is evident that about half of the candidate standard parts is not confirmed by the context analysis. Those are 42 metal rings actually forming a grid and 10 metal blocks forming the body of the product. The “42 candidate O-RINGS grouped in 42 sets” (i.e. one set for each part because they all have different diameter) are simply modeled with a single convex toroidal face, and, thus, according to the geometric requirements collected in Table 2, they are recognized as O-rings after the single parts analysis. However, the rings of the radiator are not mounted in any circular grooves, but they are welded on a radial

metal structure. As a consequence the context analysis refuses the recognition and the 42 parts are correctly discarded (Fig. 15).

Still discussing the model N.1, the “10 candidate KEYS grouped in 6 sets” (i.e. there are 2 sets made of 2 parts with same height, length, and width and other 4 parts unique in terms of dimensions) which are correctly refused by the context analysis, in reality are simple structural components having the shape of a parallelepiped (Fig. 16). In fact, they have 6 planar faces, divided in three pairs of equal and anti-parallel faces, two by two perpendicular with each other. Checking Table 2, these properties meet the requirements of keys of type B. Although, by applying the context analysis, it results that the 10 components are not keys because they are not in the corresponding seats.

The other strength of the context analysis is therefore the ability to disambiguate the assignment of a part to multiple categories that have similar shape requirements. The gearbox in model N.2 represents a remarkable example of this scenario. The single parts analysis, in fact, returns 55 components associated with more than one category. In particular, there are “13 candidate STUDS/PINS grouped in 1 set” (i.e. all the 13 parts have same dimensions) that are parts recognized both as studs and not holed pins. In fact, they have 9 faces such that 4 are planar, 2 of which are not ring, 3 are cylindrical closed not hollow, and the others are symmetric to the axis of the cylinders. By referring to Table 2,

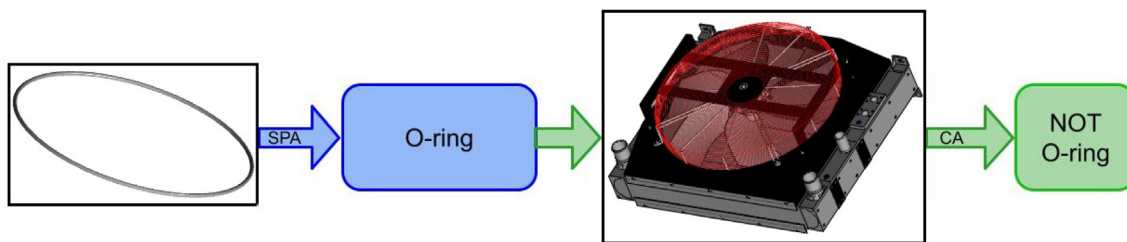


Fig. 15. Example of toroidal part recognized as O-ring by the single part analysis (SPA), but refused by the context analysis (CA).

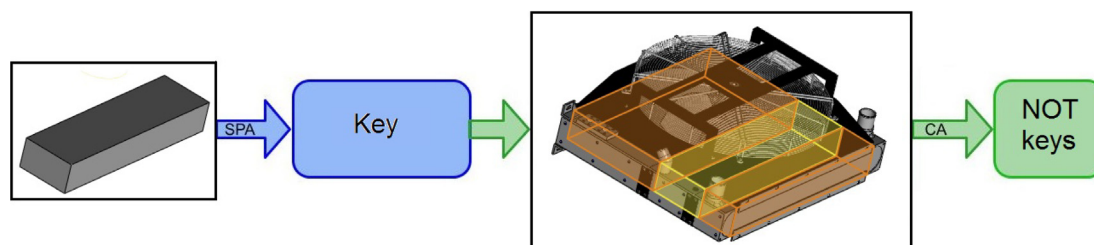
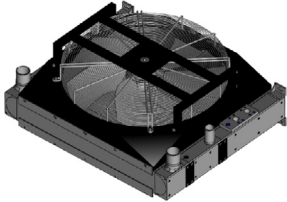
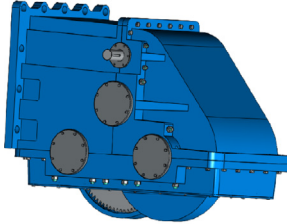
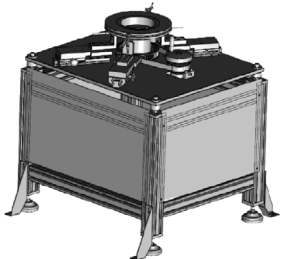
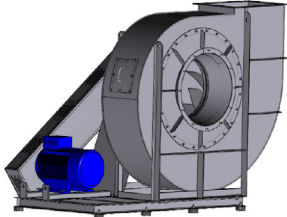
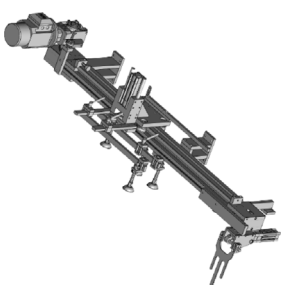
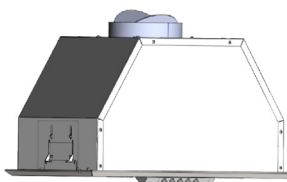


Fig. 16. Example of parallelepiped-like parts with different dimensions recognized as keys by the single part analysis (SPA), but refused by the context analysis (CA).

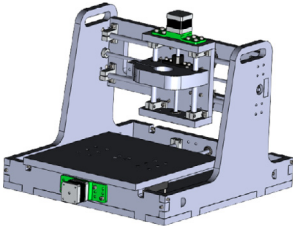
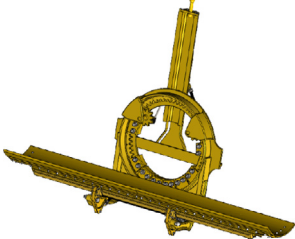
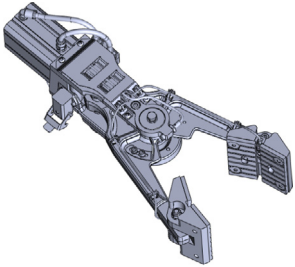

Table 4

CAD models used to validate the multi-step recognition approach. In green the parts correctly refused/assigned by the context analysis. In red the parts wrongly refused/assigned by the context analysis.

N.	CAD Model	n°parts	n°candidate standard parts	n°standard parts	Candidate standard parts refused/assigned
1	 Radiator	201	108 standard parts grouped in 55 sets of elements with same subcategory and dimensions	56 standard parts grouped in 7 sets of elements with same subcategory and dimensions	Refused: – 42 candidate O-RINGS grouped in 42 sets – 10 candidate KEYS grouped in 6 sets
2	 Gearbox	426	304 standard parts grouped in 23 sets of elements with same subcategory and dimensions	302 standard parts grouped in 22 sets of elements with same subcategory and dimensions	Refused: – 2 candidate CIRCLIPS grouped in 1 set Assigned: – 13 candidate STUDS/PINS grouped in 1 set are assigned to studs – 42 candidate CIRCLIPS/WASHERS grouped in 3 sets are assigned to WASHERS
3	 Brake disks test bed	153	44 standard parts grouped in 14 sets of elements with same subcategory and dimensions	38 standard parts grouped in 13 sets of elements with same subcategory and dimensions	Refused: – 6 candidate SCREWS grouped in 1 set
4	 Fan assembly	325	244 standard parts grouped in 31 sets of elements with same subcategory and dimensions	228 standard parts grouped in 18 sets of elements with same subcategory and dimensions	Refused: – 14 candidate KEYS grouped in 11 sets – 2 candidate CIRCLIP grouped in 2 sets
5	 Linear axis for automation	199	48 standard parts grouped in 13 sets of elements with same subcategory and dimensions	31 standard parts grouped in 7 sets of elements with same subcategory and dimensions	Refused: – 6 candidate STUDS grouped in 2 sets – 1 candidate WASHER – 2 candidate CIRCLIPS grouped in 1 set – 2 candidate WASHER grouped in 1 set – 4 candidate SCREW grouped in 1 set Assigned: – 1 candidate STUD/PIN is assigned to pin
6	 Cooker hood	106	32 standard parts grouped in 6 sets of elements with same subcategory and dimensions	31 standard parts grouped in 5 sets of elements with same subcategory and dimensions	Refused: – 1 candidate KEY

(continued on next page)

Table 4 (continued).

N.	CAD Model	n°parts	n°candidate standard parts	n°standard parts	Candidate standard parts refused/assigned
7	 Cartesian slider	237	113 standard parts grouped in 16 sets of elements with same subcategory and dimensions	104 standard parts grouped in 10 sets of elements with same subcategory and dimensions	Refused: – 9 candidate STUDS grouped in 6 sets
8	 Agricultural steel-work assembly	455	314 standard parts grouped in 18 sets of elements with same subcategory and dimensions	308 standard parts grouped in 16 sets of elements with same subcategory and dimensions	Refused: – 6 candidate STUDS grouped in 2 sets
9	 Robotic gripper	318	170 standard parts grouped in 33 sets of elements with same subcategory and dimensions	113 standard parts grouped in 29 sets of elements with same subcategory and dimensions	Refused: – 53 candidate STUDS grouped in 2 sets – 4 candidate CIRCLIP grouped in 2 sets Assigned: – 2 candidate STUDS/PINS grouped in 1 set are assigned to pin – 6 candidate CIRCLIPS/WASHERS grouped in 2 sets are assigned to WASHERS
10	 Axial reducer	306	220 standard parts grouped in 26 sets of elements with same subcategory and dimensions	173 standard parts grouped in 13 sets of elements with same subcategory and dimensions	Refused: – 6 candidate O-RING grouped in 6 sets – 16 candidate WASHER grouped in 2 sets – 14 candidate WASHER grouped in 1 set – 11 candidate CIRCLIPS grouped in 4 sets Assigned: – 1 candidate STUD/PIN is assigned to PIN

the components meet the requirements both of studs and pins. However, the context analysis correctly assigns the 13 parts to the category of studs, since contact with nuts is found (Fig. 17).

The other “42 candidate CIRCLIP/WASHERS grouped in 3 sets” (i.e. there are 3 sets of respectively 28, 8, and 6 parts having same diameters and thickness) are thus recognized both as snap rings and spring washers, due to the identical requirements for those two categories reported in Table 2. But finally the 42 components are assigned to washers because of context: they are not inserted in grooves, rather they are in contact with screws (Fig. 18).

Despite the generally good results, the context analysis in some cases fails. In fact, it can happen that standard parts resulting from the single parts analysis are wrongly rejected. However, these problems in general arise from modeling errors that are outside the tolerance threshold in the representation of features and parts (e.g. parallelism or perpendicularity between faces not detected), or from an incorrect positioning of parts in relation to the associated features (e.g. their bounding boxes never intersect within the considered tolerance), or else because some standard

parts are not even recognized by the single parts analysis and thus the context rules fail. The axial reducer in model N.10 is an example. Namely, 11 circlips are refused because misplaced with respect to the associated grooves or placed in grooves not modeled according to standards (Fig. 19(a)). Moreover, since the model is a CAD assembly created by students and thus it does not strictly follow all standards, screws are missing and consequently a set of 14 washers is refused because contacts with the respective screws is not validated. Another example is the model N.9 where, instead, a set of screws is not recognized for shape, since they do not respect some proportional requirements i.e. the ratio between the height of the head and the length of the stem is out of the standards. Consequently the washers tightened on those screws, even if they are recognized by the single parts analysis, are wrongly refused after the context evaluation (Fig. 19(b)).

Hence, wrong classifications result when modeling errors occur, but they do not totally depend on context analysis requirements, which instead turn out to be sufficiently robust. It can be broadly assessed that the standard parts that are rejected at

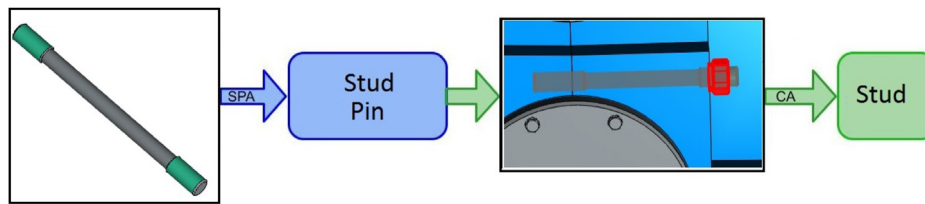


Fig. 17. Example of part recognized both as stud and pin by the single part analysis (SPA), but assigned to stud by the context analysis (CA).

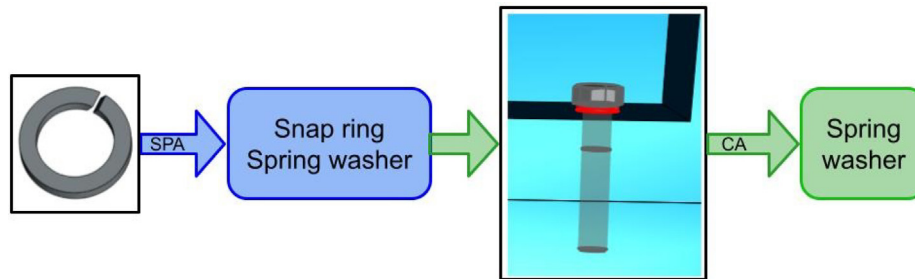


Fig. 18. Example of part recognized both as snap ring and spring washer by the single part analysis (SPA), but assigned to washers by the context analysis (CA).

this stage are lower than those that are correctly recognized, and therefore the context analysis is certainly promising and can overcome most of the issues arising from the only shape-based classification.

8. Discussion

The paper presents an innovative approach to recognize instances of some of the most used categories of mechanical parts, i.e. the standard parts, considering their shape, their conventional representation and their usage within the assembly. Distinctive features of the method are the ability to deal with industrial models, to calculate the relevant dimensions of the part and to distinguish between different sub-categories.

In order to be independent of legacy product description formats, the proposed method takes in input assembly models in STEP format. Knowing that companies and designers follow different conventions in representing standard elements [11], that, in some cases, are even omitted to maintain the model lighter, few specific assumptions and remarks are made on their representation.

First, it is required that the 3D model of the standard part is included. In fact, even if in some cases the arrangement of the features of the parts in the assembly provides hints on the presence of fixture elements, it does not allow their correct category identification. For example, the presence of threaded fasteners can be inferred by the existence of two adjacent coaxial holes (of which at least one is passing through) in two contact faces of adjacent components of the assembly, but the exact standard parts category of the fasteners is not detectable, and they may be screws as well as studs.

The other assumption made is that, even for simplified shapes, the discriminating features of the class are represented. This is the case, for instance, of the hex nut in which it is required that the part has at least the six lateral planar faces. Specifying an interval of validity for the number of faces, which is indicated as first condition among the requirements of Table 2, instead of an exact number, allows to overcome the issues induced by the different possible representations of the part, ranging from the most abstract representation, that preserves the form, to the detailed representation of the object, as available in engineering

practice. Moreover, the upper bound of the interval is also chosen to exclude not admissible candidates and, hence, make the algorithm computationally efficient.

To manage different conventions on the representation of parts, the extraction of contact information considering only the overlapping of the surfaces, even with a tolerance, is not sufficient. This is especially true for elastic parts, which can be represented in their resting or pulled position, or for threaded elements, where threads are usually represented by means of interferences with the associated hole. In these cases, in fact, intersections and/or clearances can occur, (e.g. Fig. 9), that cannot be detected. Consequently, the provided system extends the contact detection carrying out two strategies. On the one hand, cylindrical contacts are treated by assessing the ratio between the diameters of the faces is in a neighborhood of 1. On the other hand, additional reasoning on the intersection between the volumes of parts, and in particular, between the volumes of a part and of the feature associated with a seat is performed. If intersection is retrieved, then the ratios between the parameter values of the standard part and the extensions of the seat feature is evaluated. According to the type of parts, the ratios should be in given ranges to ensure the part fits the seat. Threads and deformation issues are thus overall tackled.

As for the positioning of standard parts, only considering hollow areas as features for seats is not restrictive. In fact, it is true that nuts and washers do not comply with any of these features, but their classification will be validated through the relation with other standard parts. This single type of validation is enough robust since nuts and washers fulfill their function respectively of blocking and spacing threaded fasteners only if tightened with them, and thus if a contact is detected. In addition, in mechanics it can happen that a nut is used with a different function rather than that of fastening (e.g. as spacer or resting component). In these cases the nut is not tightened on a screw/stud, and thus it is reasonably refused by the context analysis. These observations further justify the use of a multi-step approach, rather than a unique analysis of a part, that consecutively processes both the shape and the positioning inside the assembly. In fact, for some categories, e.g. washers and nuts, the single parts analysis' output is crucial to validate the recognition.

As for dimensions computation, some values are immediate to read, because intrinsic of the parts volume e.g. those regarding

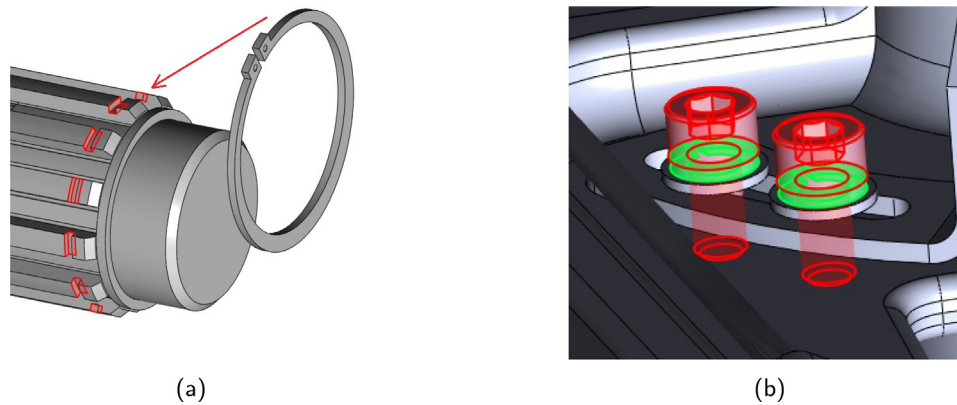


Fig. 19. Examples of standard parts not correctly recognized: (a) circlip refused for wrongly modeled groove. (b) Screws not recognized for shape and thus washer refused for context.

the extension of the parts such as length, width, and height. Other dimensions, instead concern mechanical features, such as the key size or the metric of the thread, which are not accurately represented in the model. Hence, their extraction is closely related to conventional configurations, e.g. the threads are not modeled, but their measure (i.e. the nominal diameter) corresponds to the diameter of the cylindrical face that idealizes them, while the minor diameter is that of the hole in which the fastener is inserted. However, depending on the parts representation, some dimensions may be not accessible and this is a limit of the approach. For example, if a stud is modeled as a single cylinder, the threads diameter is detected, but the length cannot be inferred since they may cover the whole stud or only a portion of it.

Then, another limit in the recognition is attributable to modeling errors that prevent the recognition of specific features. This is the case of errors affecting the shape of seats not following the mechanical standards, e.g. in a slot the side faces are not parallel even in tolerance, the lateral faces of a radial groove are modeled piecewise rather than through two complete rings (e.g. Fig. 19(a)), etc. When these types of errors occur in the models, the features associated with the seats cannot be recognized and, consequently, the context analysis would fail. The model N.2 in Table 4 is an example: the 2 candidate circlips are wrongly refused by the context analysis because they are mounted on a shaft with non standard grooves, and thus not detected by the feature recognition system. However, these problems are not related to the actual context analysis, as to the limits of the feature recognition system. Furthermore, too big displacement of parts in respect with their correct location can generate inconsistencies in the model. For example, the presence of a floating part which does not have any contact (neither by surface analysis, nor by volumes intersection) can be detected. This is a problem for the recognition, but it can be exploited to detect modeling errors and/or incomplete assemblies. Similarly, information about the rejected candidate standard parts could also provide hints of missing parts, as it was the case of the example N 10 in Table 4.

With regard to the generality of the presented method and its scalability, even if the considered classes are only a subset of the parts normalized by standard bodies, they cover the most recurrent ones within the considered types of objects, i.e. components of mechanical and electronic equipment. Anyhow the methodology can be applied to cover additional standard element categories and subcategories by defining a new recognition function for each new class. This requires the analysis of the elements of the category and the specification of rules for the identification of their salient characteristics, in terms of the number of their constituting maximal faces, mandatory faces' and edges'

types and reciprocal arrangement, and proportions between the dimensions dictated by international standards. To avoid possible overlapping or contradictory rules for standard parts, it would be convenient to connect the recognition rules to the product sector considered, i.e. it would be useless or even producing bad results applying rules related to standard components for the wood sector to mechanical parts.

Following a modular approach for rule specification and information extraction, see Fig. 1, makes possible on the one hand the reuse of the implemented functionalities for the creation of new rules; on the other hand the exploitation of information richer models than B-rep. In fact, features and liaison information, which somehow include contacts, can be made available by some CAD systems or published data models, such as [20,59,60], and could be used in our rules avoiding the data extraction step. However for completeness and generality of the method it is assumed that the data are not provided and/or available.

The use of context information, as also outlined in [11,56,60], demonstrates to be essential to avoid false positive recognition and disambiguate cases in which the shape of the parts are very similar but not their functions.

9. Conclusions

As foreseen by H. B. Voelcker [7], solid modeling has become the dominant medium for describing parts and products in the mechanical industries where 2D drafting systems have been now replaced by 3D CAD modelers. Similarly, as predicted, the availability of well-defined and complete 3D models of products opened the way for the development of vertical automation for the planning of specific processes in the product life cycle. In his vision, H. B. Voelcker pointed out the limitations in considering only forms of parts and the importance of addressing the relation between function and form as a key aspect for product designing and reasoning. Even if current CAD and PLM systems allow the annotation of product components, this process is still manual. Thus, it is still up to the designer the choice to include the information of the part meaning and function. Indeed the exploitation of formal knowledge representation and the more recent technologies mimicking the human reasoning can take advantage of functional designation of product parts for process automation.

This paper is focusing on this aspect and describes a method aimed at interpreting the functional meaning of product parts from their solid model representation. The developed multi-step part classification process applies a rule-based approach that combines engineering expertise with 3D object representation

knowledge for the specification of the distinctive conditions characterizing standard parts. In the current version of the method a rule-based approach has been adopted in which the rules are applied according to a priority dictated by practice which demonstrated that verifying the belonging of an object to a certain category according to a specific order among the categories allows to minimize the number of false positives. In the future the opportunity of using different knowledge formalization methodologies will be evaluated, for example by adopting an approach based on ontology, possibly using and extending existing ontologies [19]. Differently from most of the existing classification methods, the proposed one returns also the detailed subcategories recognition and defining important parameters. Moreover, the two-step approach allows overcoming the ambiguity of classification based solely on shape. The rule-based approach has been chosen on the one hand because of the considered engineering context, which is well structured where standard parts have to satisfy well established laws; on the other hand, the creation of the necessary training CAD assembly dataset is quite demanding. Moreover, it is easy extendable to consider other standard part classes.

The approach has been tested on industrial CAD models demonstrating its capability and the necessity of the two steps. The CAD dataset used for testing the shape-based classification step is made available at <http://standardPartRecognitionDataset.ge.imati.cnr.it> to allow a comparison with other classification methods.

The next step will be to improve and extend the presented approach, aiming at both overcoming some of the limits arisen (e.g. the presence of seats not fitting into standard features) and to enrich the set of considered parts categories and subcategories. Future work will focus on the exploitation of the obtained part recognition. Among them, the automatic decomposition of a complex assembly into smaller components according to specific characteristics can significantly contribute, for example, to assembly/disassembly planning, product redesign for additive manufacturing and product reuse.

CRediT authorship contribution statement

Brigida Bonino: Conception and design, Analysis and interpretation of data, Writing – original draft, Writing – review & editing. **Franca Giannini:** Conception and design, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Marina Monti:** Conception and design, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Roberto Raffaelli:** Conception and design, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research is carried out as part of an Industrial PhD project funded by CNR-IMATI, Italy and Hyperlean S.r.l. under the CNR-Confindustria agreement. The authors thank Anita Parodi and Paolo Barbieri for the support provided in the parts models collection and Hyperlean for the CAD models of assemblies made available.

All authors approved the version of the manuscript to be published.

References

- [1] Voelcker Herbert B, Requicha Aristides AG. Geometric modeling of mechanical parts and processes. *Computer* 1977;10(12):48–57.
- [2] Requicha Aristides AG, Voelcker Herbert B. Constructive solid geometry. *CUMINCAD*; 1977.
- [3] Requicha Aristides AG, Voelcker Herbert B. Solid modeling: Current status and research directions. *IEEE Comput Graph Appl* 1983;3(7):25–37.
- [4] Requicha Aristides AG, Voelcker Herbert B. Boolean operations in solid modeling: Boundary evaluation and merging algorithms. *Proc IEEE* 1985;73(1):30–44.
- [5] Rossignac Jaroslav R, Voelcker Herbert B. Active zones in CSG for accelerating boundary evaluation, redundancy elimination, interference detection, and shading algorithms. *ACM Trans Graph* 1988;8(1):51–87.
- [6] Shapiro Vadim, Voelcker Herb. On the role of geometry in mechanical design. *Res Eng Des* 1989;1(1):69–73.
- [7] Voelcker Herbert B. Modeling in the design process. In: *Design and analysis of integrated manufacturing system*. Washington, DC: The National Academies Press; 1988.
- [8] Voelcker Herbert B. Solid modelling: History, status, and current research directions. *Comput-Aided Geom Model* 1983;6:1.
- [9] Requicha Aristides AG, Voelcker Herbert B. Solid modeling: A historical summary and contemporary assessment. *IEEE Comput Graph Appl* 1982;2(02):9–24.
- [10] Zheng Pai, Wang Honghui, Sang Zhiqian, Zhong Ray Y, Liu Yongkui, Liu Chao, et al. Smart manufacturing systems for industry 4.0: Conceptual framework, scenarios, and future perspectives. *Front Mech Eng* 2018;13(2):137–50.
- [11] Shahwan Ahmad, Léon Jean-Claude, Foucault Gilles, Trlin Moreno, Palombi Olivier. Qualitative behavioral reasoning from components' interfaces to components' functions for DMU adaption to FE analyses. *Comput Aided Des* 2013;45(2):383–94. *Solid and Physical Modeling* 2012.
- [12] Léon Jean-Claude, Rejneri Nicolas, Debarbouille Gilles. Early assembly/disassembly simulation during a design process. In: *Proceedings of the DETC'01, ASME 2001 design engineering technical conferences and computers and information in engineering conference*. 2001, p. 1–6.
- [13] Iyer Natraj, Jayanti Subramaniam, Lou Kuiyang, Kalyanaraman Yagnanarayanan, Ramani Karthik. Three-dimensional shape searching: State-of-the-art review and future trends. *Comput Aided Des* 2005;37(5):509–30.
- [14] Carvalho LE, Von Wangenheim A. 3D object recognition and classification: A systematic literature review. *Pattern Anal Appl* 2019;22(4):1243–92.
- [15] Jayanti Subramaniam, Kalyanaraman Yagnanarayanan, Iyer Natraj, Ramani Karthik. Developing an engineering shape benchmark for CAD models. *Comput Aided Des* 2006;38(9):939–53.
- [16] Neb Alexander, Briki Iyed, Schoenhof Raoul. Development of a neural network to recognize standards and features from 3D CAD models. *Procedia CIRP* 2020;93:1429–34.
- [17] Bonino Brigida, Raffaelli Roberto, Monti Marina, Giannini Franca. A heuristic approach to detect CAD assembly clusters. *Proc CIRP* 2021;100:463–8. *31st CIRP Design Conference 2021 (CIRP Design 2021)*.
- [18] Bonino Brigida, Giannini Franca, Monti Marina, Raffaelli Roberto. Automatic assembly sequence planning for axisymmetric products. In: *Proceedings of the CAD'22 conference*. 2022, p. 334–8.
- [19] Vilmart Harold, Léon Jean-Claude, Ulliana Federico. From CAD assemblies toward knowledge-based assemblies using an intrinsic knowledge-based assembly model. *Comput-Aided Des Appl* 2018;15(1).
- [20] Hu Kai-Mo, Wang Bin, Yong Liu, Huang Jiin, Yong Jun-Hai. An extended schema and its production rule-based algorithms for assembly data exchange using IGES. *Int J Adv Manuf Technol* 2011;58:1155–70.
- [21] Raffaelli Roberto, Mandolini Marco, Germani Michele. Identification of weld beads in assemblies of B-Rep models. *Comput-Aided Des Appl* 2014;11(3):263–74.
- [22] Bonino Brigida, Giannini Franca, Monti Marina, Raffaelli Roberto. Review on the leveraging of design information in 3D CAD models for subassemblies identification. *Comput-Aided Des Appl* 2021;18(6):1247–64.
- [23] Campi Federico, Favi Claudio, Germani Michele, Mandolini Marco. CAD-integrated design for manufacturing and assembly in mechanical design. *Int J Comput Integr Manuf* 2022;35(3):282–325.
- [24] Kretschmer Ralf, Pfouga Alain, Rulhoff Stefan, Stjepandić J. Knowledge-based design for assembly in agile manufacturing by using data mining methods. *Adv Eng Inform* 2017;33:285–99.
- [25] Tasalloti Hamed, Eskelinen Harri, Kah Paul, Martikainen Jukka. An integrated DFMA-PDM model for the design and analysis of challenging similar and dissimilar welds. *Mater Des* 2016;89:421–31.
- [26] Hadj Riadh Ben, Belhadj Imen, Trigui Moez, Aifaoui Nizar. Assembly sequences plan generation using features simplification. *Adv Eng Softw* 2018;119:1–11.
- [27] Wu Meiping, Zhao Yi, Wang Chenxin. Knowledge-based approach to assembly sequence planning for wind-driven generator. *Math Probl Eng* 2013;2013.

- [28] Pan Wanbin, Wang Yigang, Chen Xiao-Diao. Domain knowledge based non-linear assembly sequence planning for furniture products. *J Manuf Syst* 2018;49:226–44.
- [29] Qiao Lihong, Qie Yifan, Zhu Zuwei, Zhu Yixin, uz Zaman Uzair Khaleeq, Anwer Nabil. An ontology-based modelling and reasoning framework for assembly sequence planning. *Int J Adv Manuf Technol* 2018;94(9):4187–97.
- [30] Shi Xiaolin, Tian Xitian, Wang Gangfeng, Zhao Dongping, Zhang Min. Semantic-based subassembly identification considering non-geometric structure attributes and assembly process factors. *Int J Adv Manuf Technol* 2020;110(1):439–55.
- [31] Yin ZhouPing, Ding Hang, Li HanXiong, Xiong YouLun. A connector-based hierarchical approach to assembly sequence planning for mechanical assemblies. *Comput Aided Des* 2003;35:37–56.
- [32] Fang HC, Ong SK, Nee AYC. Product remanufacturability assessment based on design information. *Proc CIRP* 2014;15:195–200, 21st CIRP Conference on Life Cycle Engineering.
- [33] Li JingRong, Wang QingHui, Huang Ping, Shen HongZhou. A novel connector-knowledge-based approach for disassembly precedence constraint generation. *Int J Adv Manuf Technol* 2010;49(1):293–304.
- [34] Mandolini Marco, Favi Claudio, Germani Michele, Marconi Marco. Time-based disassembly method: How to assess the best disassembly sequence and time of target components in complex products. *Int J Adv Manuf Technol* 2018;95:409–30.
- [35] Zhang Chao, Zhou Guanghui, Lu Qi, Chang Fengtian. Generating significant subassemblies from 3D assembly models for design reuse. *Int J Prod Res* 2018;56(14):4744–61.
- [36] Fradi Ahmed, Louhichi Borhen, Mahjoub Mohamed Ali, Eynard Benoit. A new approach for reusable 3D CAD objects detection, by similarity calculation based on Bayesian network models (BNM). *Int J Comput Integr Manuf* 2021;1–20.
- [37] Huang Rui, Zhang Shusheng, Bai Xiaoliang, Xu Changhong, Huang Bo. An effective subpart retrieval approach of 3D CAD models for manufacturing process reuse. *Comput Ind* 2015;67:38–53.
- [38] Qin Feiwei, Gao Shuming, Yang Xiaoling, Li Ming, Bai Jing. An ontology-based semantic retrieval approach for heterogeneous 3D CAD models. *Adv Eng Inform* 2016;30(4):751–68.
- [39] Biasotti Silvia, Marini Simone, Spagnuolo Michela, Falcidieno Bianca. Subpart correspondence by structural descriptors of 3D shapes. *Comput Aided Des* 2006;38(9):1002–19, Shape Similarity Detection and Search for CAD/CAE Applications.
- [40] Jayanti Subramaniam, Kalyanaraman Yagnanarayanan, Ramani Karthik. Shape-based clustering for 3D CAD objects: A comparative study of effectiveness. *Comput Aided Des* 2009;41(12):999–1007.
- [41] Liu Zhen-Bao, Bu Shu-Hui, Zhou Kun, Gao Shu-Ming, Han Jun-Wei, Wu Jun. A survey on partial retrieval of 3D shapes. *J Comput Sci Tech* 2013;28:836–51.
- [42] Hu Kai-Mo, Wang Bin, Yong Jun-Hai, Paul Jean-Claude. Relaxed lightweight assembly retrieval using vector space model. *Comput Aided Des* 2013;45(3):739–50.
- [43] Lupinetti Katia, Pernot Jean-Philippe, Monti Marina, Giannini Franca. Content-based CAD assembly model retrieval: Survey and future challenges. *Comput Aided Des* 2019;113:62–81.
- [44] Li Min, Zhang YF, Fuh JYH, Qiu ZM. Toward effective mechanical design reuse: CAD model retrieval based on general and partial shapes. *J Mech Des* 2009;131(12).
- [45] Ip Cheuk Yiu, Regli William C, Sieger Leonard, Shokoufandeh Ali. Automated learning of model classifications. In: *Proceedings of the eighth ACM symposium on solid modeling and applications*. 2003, p. 322–7.
- [46] Pernot Jean-Philippe, Giannini Franca, Petton Cédric. Thin part identification for CAD model classification. *Eng Comput* 2015;32(1):62–85.
- [47] Ip Cheuk Yiu, Regli William C. A 3D object classifier for discriminating manufacturing processes. *Comput Graph* 2006;30(6):903–16.
- [48] Cicek Adem, Gülesin Mahmut. A part recognition based computer aided assembly system. *Comput Ind* 2007;58(8–9):733–46.
- [49] Manda Bharadwaj, Bhaskare Pranjali, Muthuganapathy Ramanathan. A convolutional neural network approach to the classification of engineering models. *IEEE Access* 2021;9:22711–23.
- [50] Wei Liu, Yuanjun He. Representation and retrieval of 3D CAD models in parts library. *Int J Adv Manuf Technol* 2008;36(9):950–8.
- [51] Rucco Matteo, Giannini Franca, Lupinetti Katia, Monti Marina. A methodology for part classification with supervised machine learning. *AI EDAM* 2019;33(1):100–13.
- [52] Bonino Brigida, Giannini Franca, Monti Marina, Raffaelli Roberto. Enhancing product semantics understanding through automatic part type recognition in CAD assembly models. *Comput-Aided Des Appl* 2022;19(5):896–912.
- [53] Bloch Tanya, Sacks Rafael. Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. *Autom Constr* 2018;91:256–72.
- [54] Jian Chengfeng, Liang Lingming, Qiu Keyi, Zhang Meiyu. An improved memory networks based product model classification method. *Int J Comput Integr Manuf* 2021;34(3):293–306.
- [55] Foucault Gilles, Léon Jean-Claude. Enriching assembly CAD models with functional and mechanical informations to ease CAE. In: *International design engineering technical conferences and computers and information in engineering conference*, vol. 44113. 2010, p. 341–51.
- [56] Lupinetti Katia, Giannini Franca, Monti Marina, Pernot Jean-Philippe. Identification of functional components in mechanical assemblies. *Proc CIRP* 2017;60:542–7.
- [57] Lupinetti Katia, Giannini Franca, Monti Marina, Rucco Matteo, Pernot Jean-Philippe. Identification of functional sets in mechanical assembly models. In: *International conference on innovative design and manufacturing*. 2017, p. 1–6.
- [58] Di Stefano Paolo, Bianconi Francesco, Di Angelo Luca. An approach for feature semantics recognition in geometric models. *Comput Aided Des* 2004;36(10):993–1009.
- [59] Rachuri Sudarsan, Baysal Mehmet, Roy Utpal, Fofou Sebti, Bock Conrad, Fenves Steven J, et al. Information models for product representation: Core and assembly models. *Int J Prod Dev* 2005;2(3):207–35.
- [60] Swain Abinash Kumar, Sen Dibakar, Gurumoorthy B. Extended liaison as an interface between product and process model in assembly. *Robot Comput-Integr Manuf* 2014;30(5):527–45.