Journal of Advanced Zoology
ISSN: 0253-7214
Volume 45 Issue 4 Year 2024 Page 106-116

An Automatic Bioinformatic System For Identification And Classification Of Hymenoptera (The Apidae Family)

Aouatef CHAIB[1*], Malak Intissar DJENDLI[2], Youssra LEKBIR[2]

[1]Laboratory of Biosystematics and Ecology of Arthropos, Department of Animal Biology, Faculty of Natural and Life Sciences, Mentouri Brothers University of Constantine 1, 25000, Constantine, DZ-Algeria., [2]Independent researcher, Constantine, Algeria.*

***Corresponding Author:** Aouatef CHAIB*
E-mail: aouatef.chaib@umc.edu.dz

Article History

Received: 18 January 2024
Accepted: February 25, 2024
Publication date: April 2, 2024

Abstract

In this paper we illustrate our bioinformatics system destined for an automatic insect classification in different taxonomic categories corresponding to different levels of precision according to their morphological characteristics such as: size; color; antennae; wings; mouthpart; etc. Our main goal is to automate the research; classification and to create an exhaustive database of insects in order to create a common directory of insects to save significant research time and facilitate classification.

Our system is considered as a bioinformatics system, it uses artificial intelligence techniques to create an inference engine able to automatically recognize and classify a new insect. It also uses database creation techniques to create the knowledge and fact bases necessary for the proper functioning of the inference engine.

Due to its interactive graphic interface, and by guiding its user, step by step, the system starts with the introduction of the morphological characteristics of the insect and the result is to automatically classify the insect in the database in the appropriate species.

Our classification is focused on Hymenoptera precisely on the Apidae family, we have given a detailed classification of the Apidae family with its three subfamilies: Apinae, Xylocopinae and Nomadini.

Our system is implemented completely in JAVA programming language using phpMyadmin as a database management system.

***Keywords:** Bioinformatic; Automatique classification; Hymenoptera, Apidae family, Data bases.*

INTRODUCTION

The Scientific or Systematic Classification of Species Coming from the classification of living organisms is generally based on morphological characteristics. In entomology, classifying an insect consists of placing it in different hierarchical taxonomic categories corresponding to different levels of precision: phylum, class, order, family, genus and species [1][2]. By dint of the use of computer resources and techniques,

classification has become an automatic classification or supervised classification. It consists of assigning a class or category to each object (or individual) to be classified, based on statistical data. It commonly uses machine learning and is widely used in pattern recognition. This marriage between biology and computer science is called Bioinformatics.

Bioinformatics [3] has become one of the main methods of analysis in the daily life of biological researchers. It is a so-called *in silico* approach which complements *in vivo* and *in vitro* studies. The challenge of Bioinformatics is twofold with on the one hand the development of methods of acquisition, control and analysis of data, and on the other hand the transition from the level of data analysis to that of knowledge [4][5].

In entomology, the application of bioinformatics in entomology is well evolving, which allows the entomologist to express and develop software tools to understand biological data. Several works have been introduced; a comparative study of some works is given in the Table 1.

FPM (fly picture Measurement)[4]: is a software tool for measuring the size of remarkable segments and the mean gray level of the wing was designed and applied to the wing of tsetse fly (Diptera: Glossinida). It makes it possible to quickly carry out a set of measurements in a semi-automatic way. On the image, the user selects with his mouse eight remarkable points, located intersections of ribs. This software gives access to a set of parameters that cannot be measured visually or with very uncertain precision. Then, the results are reproducible and usable for reliable statistical analyses.

Xper3 [6]: It is dedicated to the management of structured taxonomic descriptions and their analysis. It contains several tools for analyzing the consistency of descriptive data. It is equipped with a module for generating interactive identification keys with free access from the databases stored in the system, Xper3 also offers the ability to share data with third-party users, making collaborative management of descriptive data possible.

Forenseek [7] : Is a computer tool for decision support, dedicated to the realization of simulations and expertise in forensic entomology. The core of this program is a model of behavior and development of necrophagous Diptera larvae based on a multi-agent system (MAS). This process makes it possible to model the behavior of complex groups from simple individual data, and is therefore particularly well suited to the study of insect populations. The results of this model are intended to be automatically processed to extract information on the relevance of each simulation, and thus to estimate the post-mortem interval (PMI) in an automated and precise manner.

EgBugs [8]: The authors aim, in the construction of the database, to be an easy-to-use and efficient tool for assimilating such data. By extracting information from hundreds of entomological sources in one program, adding various search and report functions, it is hoped that EgBugs will significantly reduce the time its users have to spend on analog data mining.

Table 1. Comparative study of the different works in entomology which uses bioinformatics

Project	**Data bases**	**Morphological characteristics**	**Insect recognition**	**Décision**	**Domain**
FPM	-	Taille insect L'aile	Semi automatic	-	Medical domain
Xper3	++	/	automatic	-	Study and reseach
Forenseek	-	/	automatic	++	Forensic medicine Insect populations
EgBugs	+	+	automatic	-	agriculture
Our project	++	All morphological characteristics	automatic	++	Apiculture Food Insect classification

Our system is intended for different types of users such as entomologist researchers, school public (students and teachers), field professionals, volunteers... it provides a simple and user-friendly interface and ease of use (modification, addition, storage, research…).

The main objective of this project is to provide entomologists with an exhaustive database of insects and create a common directory to automate research and recognition of new species for use in other automatic recognition and classification systems of insects. Searching for insect and species distribution data quickly on our system saves significant time from searching in libraries and on the Internet.

MATERIAL AND METHODS

Study population and area

In our project we have choosing the Hymenoptera order [9] for several reasons [10][11][12]:

- The Hymenoptera are among the most important orders of insects.
- They constituting one of the largest groups of insects, is represented by some 120,000 species,
- Hymenoptera are best known for their economic role (honey production), or beneficial (nectar lickers contribute to the cross-pollination of plants), or harmful (ants).
- Hymenoptera are very essential in the natural balance

Our classification of Hymenoptera [9] is focused precisely on the Apidae family, we have given a detailed classification of the Apidae family with its three subfamilies: Apinae, Xylocopinae and Nomadini.
The body morphology of Apidae is particularly well adapted to pollination, it is most often characterized by [13]: very abundant hair, the hairs (or bristles) are feathery (characteristic of apoids) which facilitates the adhesion of the pollen grains when the insect visits the flowers and allows the bee to transport the pollen.
Our study is a synthesis of several researches on hymenoptaria [14][15][16][17].

The computer tools

JAVA language

In order to create the interface allowing users to manipulate the prototype, we chose the Java language. This choice was motivated by the following reasons:

- ✓ Java ensures total independence of applications vis-à-vis the execution environment: i.e. any machine supporting Java is able to execute a program without any adaptation.
- ✓ Simplifies access to data bases. Environnement de développement

Regarding the development environment, we chose Net Beans 12.1 which is an integrated development environment (IDE) for the development of object-oriented applications.

Php Myadmin

Is a management Web application for the MySQL and Maria DB database management systems, made mainly in PHP and distributed under the GNU GPL license. This is one of the most famous interfaces for managing a MySQL database on a PHP server. Many hosts, both free and paid, offer it, which saves the user from having to install it.

Proposed system

Our project named Automatic Hymenoptera Classification System (AHCS) is an automatic classification system for Hymenoptera (Apidae family) our classification is according to the morphological characteristics of the insect such as: size; color; antennae; wings; mouthpart, etc. The system contains a recognition and classification engine and is equipped with two databases: Hymenoptera Data base and Facts Data base. A general description of our system is summarized in the Fig.1.

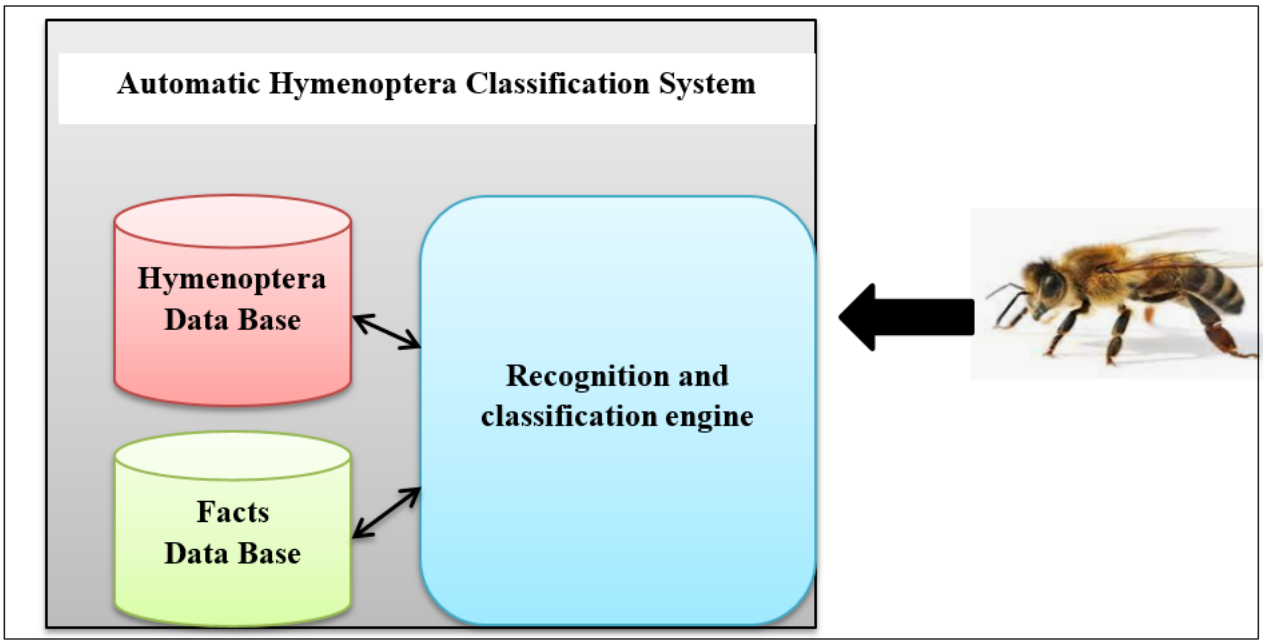

Fig.1. General description of the system

The system allows to:
- ✓ Add/ Identify an insect.
- ✓ Save morphological characteristics of insects.
- ✓ Store and edit insects
- ✓ Search insects by their morphological characteristics.
- ✓ Automatically recognize and classify the species of a new insect
- ✓ Find all the information for a given insect

AHCS is intended for different types of users such as entomologist researchers; school public (students and teachers); field professionals; volunteers; etc. It provides a simple and user-friendly interface and ease of modification; addition; storage; research; etc.
The main objective of AHCS is to provide entomologists with an exhaustive database of Hymenoptera and create a common directory to automate research and recognition of new species for use in other automatic recognition and classification systems of insects. Searching for insect and species distribution data quickly on our system saves significant time from searching in libraries and on the Internet.

RESULTS AND DISCUSSION

Hymenoptera classification

Our classification of the third sub family Apinae, Xylocopinae and Nomadini; we have dividing our classification into several tribes [18][19]. We have also defined the common name of insects of the same tribes. The morphological characteristics that we used in our classification are the most used in automatic insect recognition systems such as: the insect size, the body of the insect which itself has several morphological characteristics: head; thorax; Clypeus; Mandible; Antenna; Abdomen; tergites; color and paw. We have also added distribution to our classification to make it easier for the searcher to target his search area exactly. Tables 2- 4 represent a part of our classification of the three subfamilies Apinae; Xylocopinae and Nomadini in this order.

Table 2. A part of Apinae classification

		Common name	Insect size	Body	Color	Paw	Distribution
Anthophorini Tribe	**Anthophora Albigena**		10 mm.	**Head:** Black White Wrasse. **Abdomen:** black bristling with hair. **Thorax:** black and bristling with hair. **Wings:** transparent black nerves a little smoky. **Hairiness:** Reddish-white hair (cluster of whitish hair).	Black	Black bristles with hair.	East and West of Algeria Around Oran and Algiers Lyon and Sicily.
	Anthophora Robusta	Robust anthophore.	17-19 mm.	**Head:** Short, regular sculpted coat **Abdomen:** T1 _2 creamy white regular short erect coat. **Thorax:** regular short straw yellow coat with black bristles.	Black	Straw yellow to golden paw coat	Most countries bordering the Mediterranean Constantine; Eastern Algeria.

Table 3. A part of Xylocopinae classification

		Common name	Insect size	Body	Color	Paw	Distribution
Xylocopini Tribe	**Xylocopa pubescens**	Carpenter bee.	Large and shiny females the males are smaller 13-30 mm.	**Clypeus:** Male: entirely dark. **female:** absence of punctuate longitudinal line in the middle of the clypeus. **Hairiness:** last abdominal segments hairier than the rest of the body, the hairs longer and denser.	**Females:** black with yellow markings on the head. **Males:** narrow head and yellow pubescence that covers their entire body.	the base of the posterior femora simple the apical process of the posterior tibia more developed .	North Africa (Morocco, Tunisia, Algeria, Libya, Egypt).

	Xyloco paustulata	Carpenter bee.		**Tergite:** The last tergite is bifid **Hairiness:** last abdominal segments more hairy than the rest of the body, the hairs longer and denser. the posterior tibiae and metatarsals The setae of the face are rather gray.	**Male:** black face with rich gray or black pubescence	the hind tibiae and metatarsals completely furnished with a rather shaggy black pubescence.	Algeria, (Hoggar¸ Oued¸ Tekouiat¸ In Amguel¸ Tamanrasset) Libya, Egypt, Mauritania, Gambia, Mali, Burkina-Faso, Niger, Nigeria, Cameroon.

Table 4. A part of Nomadini classification

		Common name	**Insect size**	**Body**	**Color**	**Paw**	**Distribution**
nomadini Tribe	**Nomada Rubiginosa**		**Male :** 7_8 mm. **female :** 6_8 mm.	**Clypeus:** Enlarged yellow in the middle. **Mandible:** brown. **Antenna:** not clavate. **Abdomen:** red color. **Tergite:** very distinct.		Posterior metatarsus slightly obscured or red.	Turkey France Italy Spain North Africa (Tunisia, Algeria.
	Nomada Glaucopis		**Male :** 6_8 mm **female :** 6_7,5 mm.	**Abdomen:** red. **Head and Thorax:** black. Clypeus **Mandible:** yellow.		Prototarsus posterior short thick posterior tarsus ringed with brown.	France Italy Spain Portugal North Africa Sahara western Algeria.
	Nomada Dolosa		6,5 mm.	**Head:** yellowish gray. **Thorax:** hairy. **Abdomen:** elongated		Reddish color.	France; Italy Spain Portugal Sahara western Algeria.
	Nomada Brevis		6mm.	**Abdomen:** longer than the thorax. **Head and Thorax:** black. **Wings and Rib:** dark.		Tibials posterior armed externally at the apex of 3 short spines obtuse blades	North Africa Portugal Slovenia Greece Israel Algeria (Biskra)
nomadini Tribe	**Nomada Numida**		14 _16 mm.	**Head And Thorax:** covered with brown hair		Tibial apices posterior lying with two thorns curved.	Portugal Slovenia Greece Israel North Africa Algeria (the surroundings of Oran).

Internal presentation of the system

Our system works according to three different complementary modules which are: Recognition module; Classification module and Communication module, Fig.2 represents the detailed internal architecture of our system.

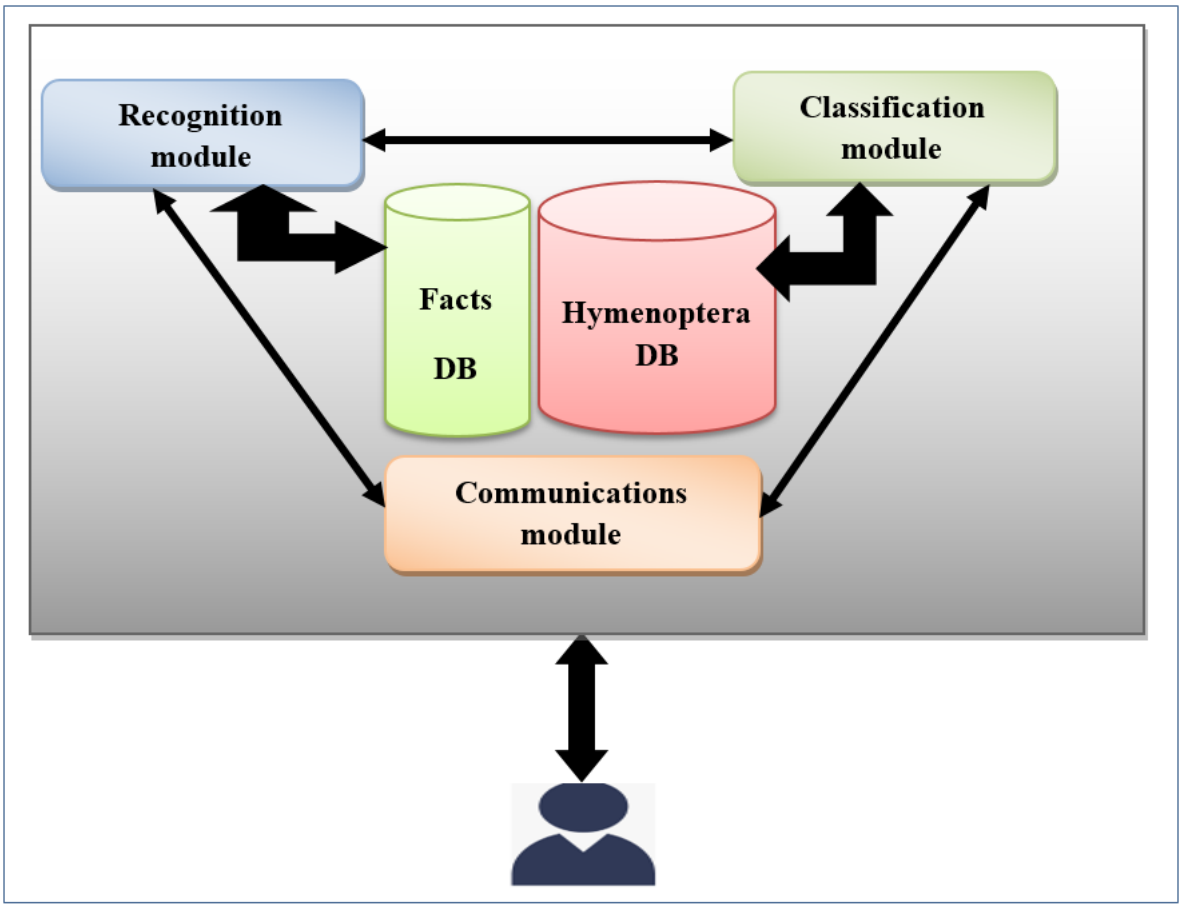

Fig.2. Internal architecture of the system

Communications module

- Receives the morphological characteristics of the insect.
- Send the morphological characteristics of the new insect to the Recognition module.
- Receives the class of the new insect from the Recognition module.
- Sends the class of the new decision module insect to the user.

Recognition Module

- Receives the morphological characteristics of the insect from Communication module .
- Interrogates Fact Data base.
- Know the subfamily and the species.
- Send the family and the species to the Classification module.

Classification Module

- Receives the morphological characteristics of the new insect from the Communication module.
- Receives the new insect subfamily from the Recognition module.
- Creates a new record of the new insect in the adequate table of Hymenoptera database.

Hymenoptera data base

Our system has an acquaintance database implemented at its level, it is a data structure that is updated cyclically; it contains all the insects and information (morphological characteristics) of the Apidae family. Our database contains three large tables which are: the Apinae table; the Xylocopinae table and the Nomadini table.

Facts data base

This database contains the facts and reasoning rules of our system; its state evolves during the expertise (working memory). This is the result of a query from an expert in the field. This expertise is represented in the form of type rules:

If (Condition1; Condition2; …. Condition n) then (Action1; Action 2; ….; Action n)

System operation

In the Fig.3 we present the operating diagram of our system with the interactions between its modules and the databases

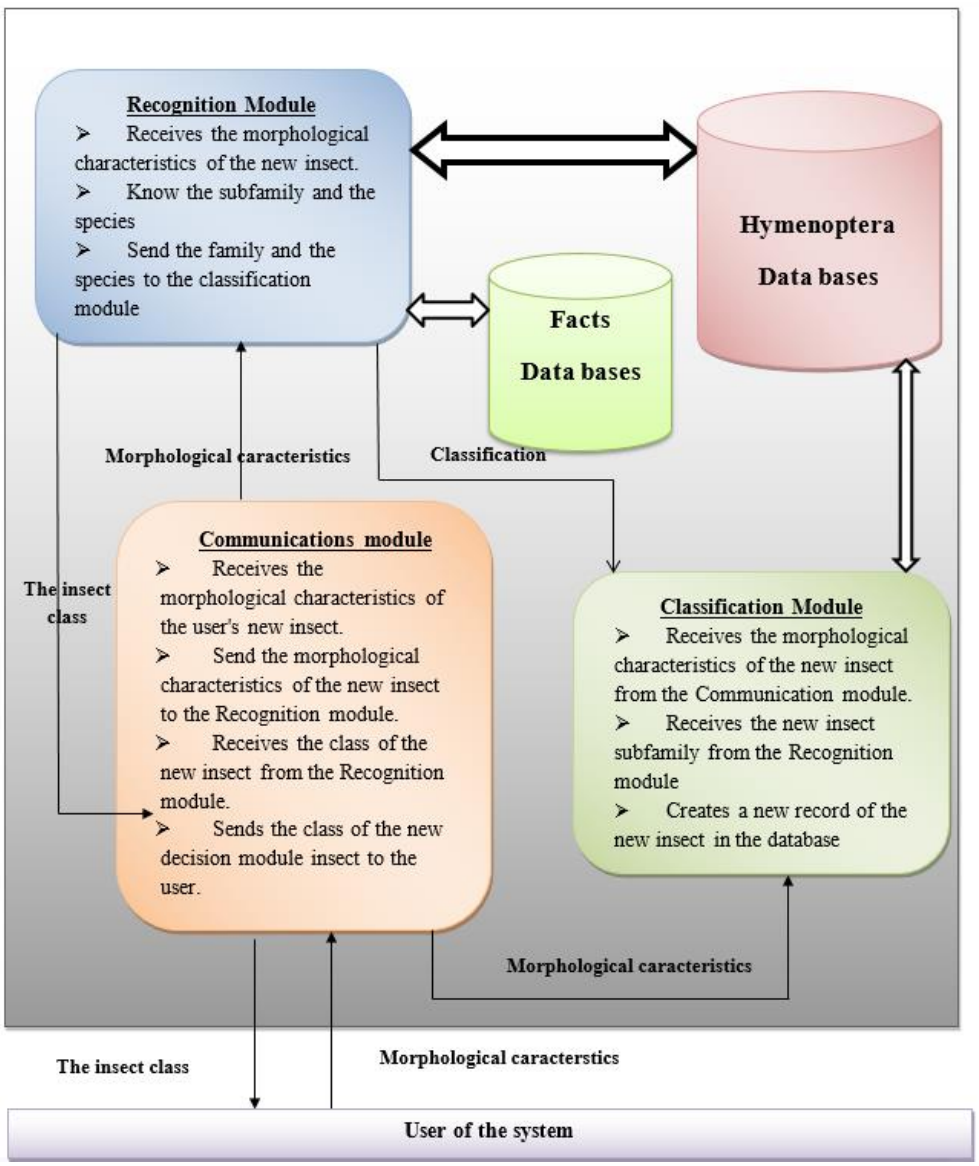

Fig. 3. Functional diagram of the Automatic Hymenoptera Classification

System execution

The home interface of our system is illustrated in Fig.4. This latter is displayed as soon as we launch the execution of the system

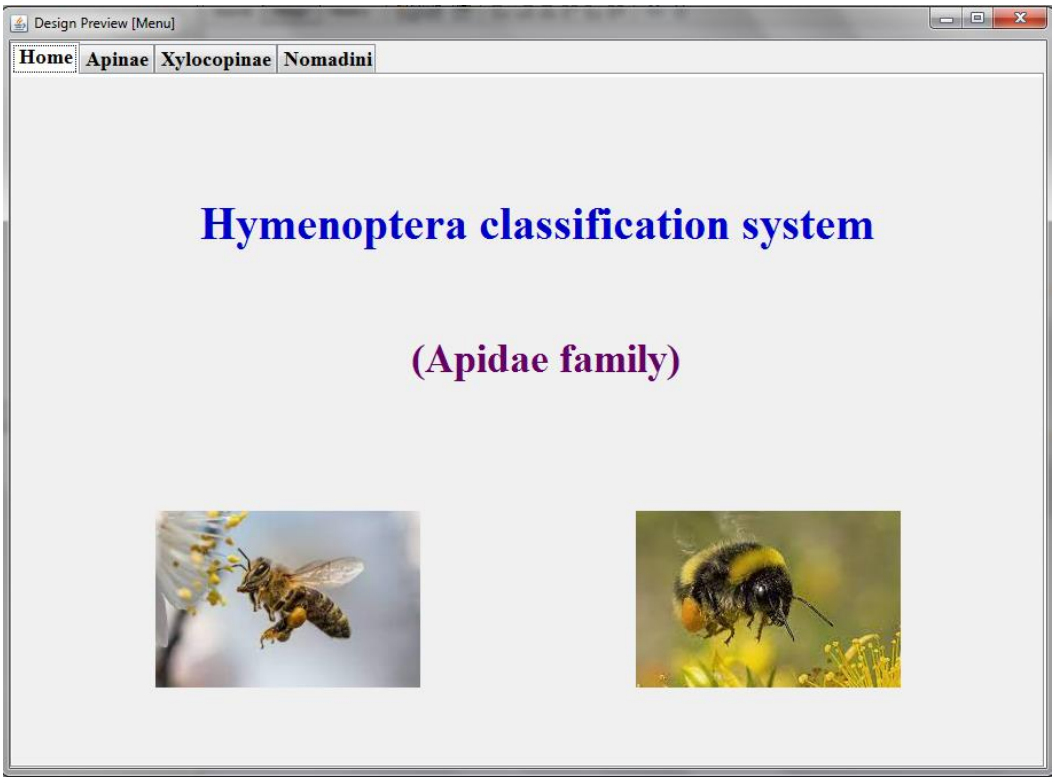

Fig.4. Home interface.

The home interface of the proposed system has three panes: Apidae pane, Xylocopinae pane and Nomadini pane. Each pane allows a classification of it sub-family according to its own morphological characteristics. Each pane is a set of text fields that allow user to enter the obtained morphological characteristics; this latter are necessary in the cognition and the classification of the insect Figs.5-7. The pane also contains two buttons:

***The Add button*:** in the case of a new insect, it allows the user to create an entry for him in the Hymenoptera database.

***The Search button*:** allows the user to search the detail of a the insect already registered in the system.

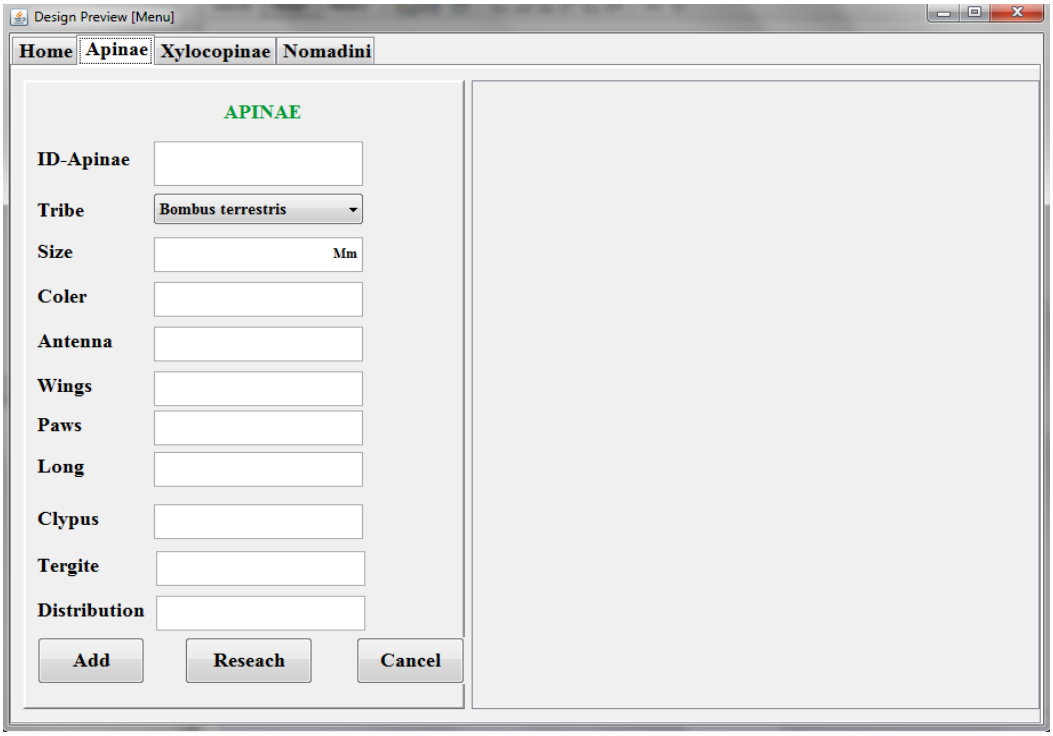

Fig.5. Apinae Pane.

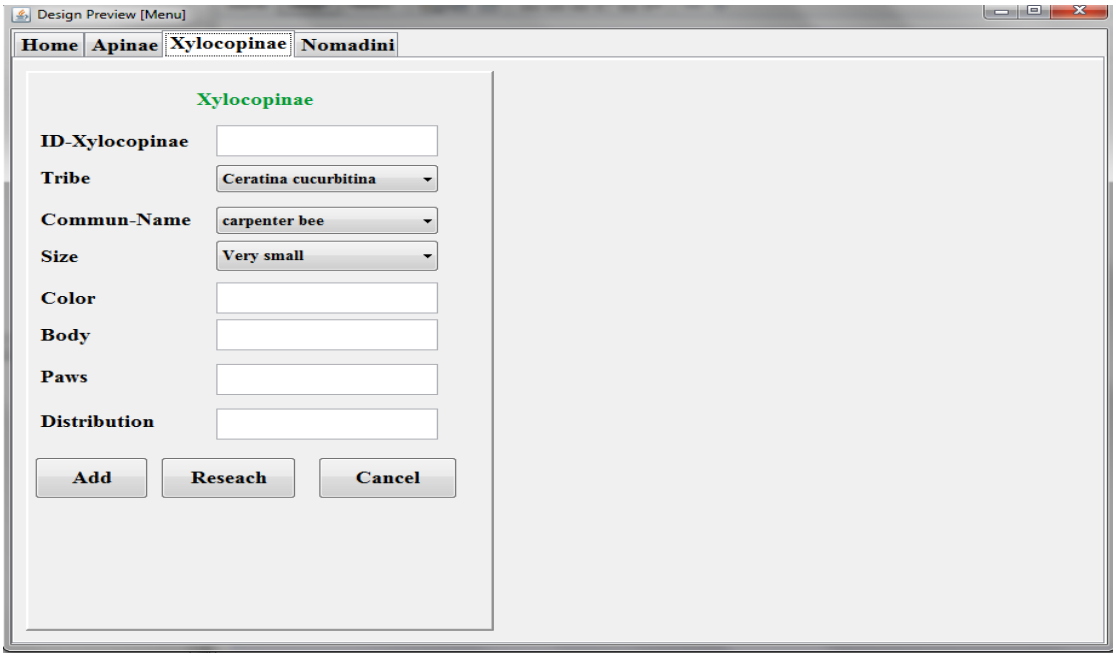

Fig.6. Xylocopinae Pane.

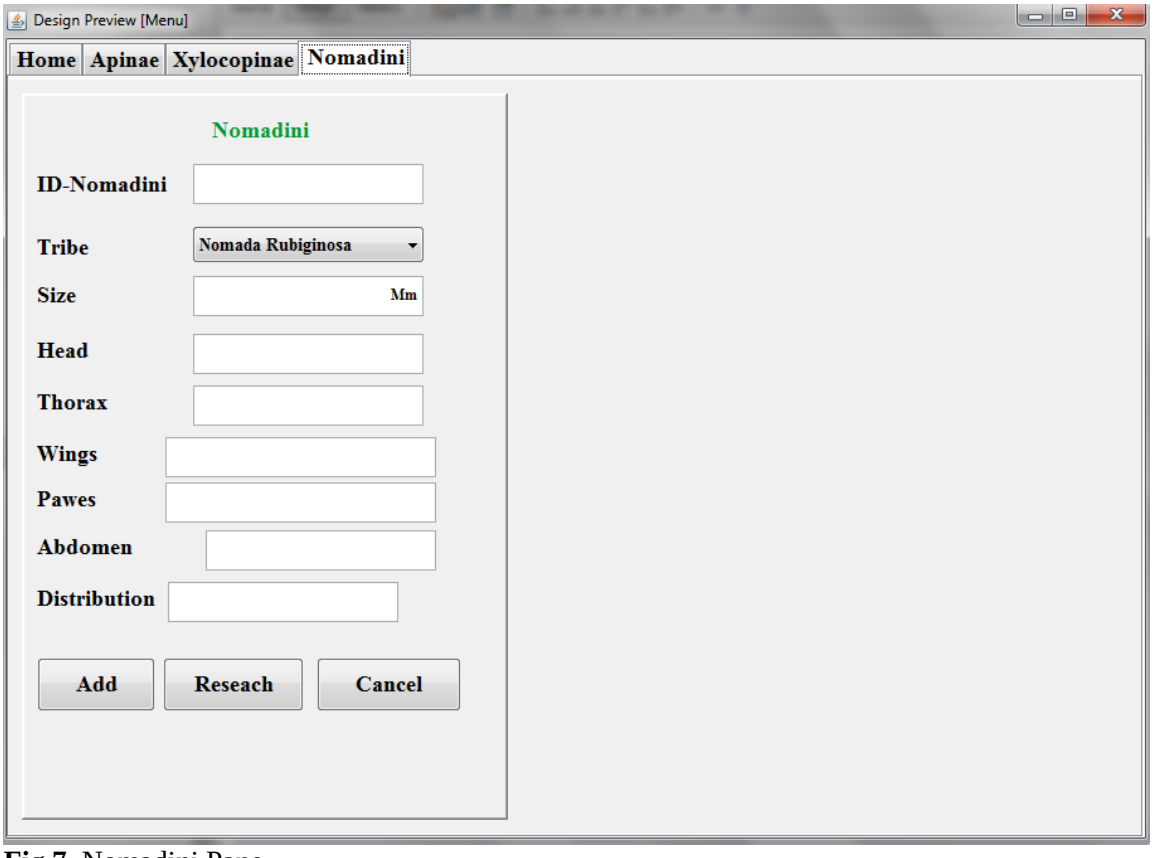

Fig.7. Nomadini Pane.

CONCLUSION

Automatic Hymenoptera Classification System (AHCS) allows automatic recognition and classification of Hymenoptera quickly and accurately due to an inference engine and databases designed for this purpose. The identification of the species is based on the morphological characteristics of the insect. It represents an effective and reliable tool for the identification and classification of insects; moreover it allows also to create an electronic directory of the Apidae family.

AHCS has two databases for a better implementation of the inference engine; one to keep the morphological characters of insect and the other for the expertise of our expert system. So AHCS is a database manager;

works are in progress in order to enrich its base of facts by the maximum possible of expertise to generalize its use for the maximum insect species.
An improvement on identification method is in progress; using the identification key. Integration of a new component in the system of automatically taking morphological characteristics from a photo taken by the user using image processing techniques.

REFERENCES

1. Wen, C., & Guyer, D. (2012). Image-based orchard insect automated identification and classification method. *Computers and electronics in agriculture*, *89*, 110-115.
2. Martineau, M., Conte, D., Raveaux, R., Arnault, I., Munier, D., & Venturini, G. (2017). A survey on image-based insect classification. *Pattern Recognition*, *65*, 273-284.
3. BOUADJILA, L., & TALEB, Z. *Le portail NCBI: base de données bioinformatique clé en biotechnologies* (Doctoral dissertation, Université Mohamed BOUDIAF de M'Sila).
4. Borne, F., Petiteau, L., Geoffroy, B., De La Rocque, S., & Cuisance, D. (1999). Fly Picture Measurement, un nouvel outil informatique pour l'étude des glossines. *Revue d'élevage et de médecine vétérinaire des pays tropicaux*, *52*(1), 19-21.
5. Rialle, S. (2010). *Méthodologie et outils bioinformatiques d'aide à la conception de systèmes biologiques synthétiques pour de nouveaux diagnostics en santé humaine* (Doctoral dissertation, Montpellier 2).
6. Delaunay, M. (2019). *Comment faciliter l'identification de l'entomofaune?: Construction, évaluation et amélioration de clés d'identification numériques* (Doctoral dissertation, Museum national d'histoire naturelle-MNHN PARIS).
7. Charabidze, D., Morvan, G., Dupont, D., Gosset, D., & Bourel, B. (2008, January). ForenSeek: un programme de simulation du développement des insectes nécrophages dédié à l'entomologie médico-légale. In *Annales de la Société Entomologique de France* (Vol. 44, No. 3, pp. 385-392). Taylor & Francis Group.
8. Buckland, P. I., Buckland, P. C., Panagiotakopulu, E., & Sadler, J. P. (2004). A database for Egyptian entomology. *Bulletin of the Entomological Society of Egypt*, *81*, 1-16.
9. Dufour, M. L., & Perris, É. (1840). Mémoire sur les Insectes Hyménoptères qui nichent dans l'intérieur des tiges sèches de la Ronce. *Ann. Soc. Ent. France*, *9*, 3.
10. Delvare, G., & Aberlenc, H. P. (1989). *Les insectes d'Afrique et d'Amérique tropicale: clés pour la reconnaissance des familles*. Editions Quae.
11. Noblecourt, T. (2018). Liste systématique des Hyménoptères Symphytes de France (2018).
12. Blanchard, E. (1845). *Histoire des insectes: traitant de leurs moeurs et de leurs métamorphoses en général et comprenant une nouvelle classification fondée sur leurs rapports naturels* (Vol. 2). Firmin Didot Frères.
13. Maghni, N., & Louadi, K. (2017). *Biogéographie des apidae (hymenoptera; apoidea) d'algérie et monographie des espèces d'eucerini et anthophorini dans la région des aurès* (Doctoral dissertation, Université Frères Mentouri-Constantine 1).
14. Smit, J. Identification Key to the European Species of the bee genus Nomada Scopoli 1770 (Hymenoptera: Apidae).
15. Latreille, P. A. (1831). *Cours d'entomologie, ou de l'histoire naturelle des crustacés, des arachnides, des myriapodes et des insectes: à l'usage des élèves de l'école du Muséum d'histoire naturelle*. Roret.
16. Pecault, F. (2002). *L'envenimation par les hyménoptères* (Doctoral dissertation). Reymond, N. (2004). *Bioinformatique des puces à ADN et application à l'analyse du transcriptome de Buchnera aphidicola* (Doctoral dissertation, INSA de Lyon).
17. Maghni, N., Louadi, K., Ortiz-Sánchez, F. J., & Rasmont, P. (2017, January). Les Anthophores de la région des Aurès (nord-est de l'Algérie)(Hymenoptera: Apidae: Anthophorini). In *Annales de la Société entomologique de France (NS)* (Vol. 53, No. 1, pp. 55-73). Taylor & Francis.
18. Michener, C. D. (1996). A review of the genera of Brachynomadini and a new South American genus (Hymenoptera, Apidae, Nomadinae). *Journal of the Kansas Entomological Society*, 87-96.
19. Darchen, R., & Delage-Darchen, B. (1975). Contribution a l'Étude d'une Abeille du Mexique Melipona Beecheii B.(Hymenoptère: Apide). Le Déterminisme des Castes Chez les Mélipones. *Apidologie*, *6*(4), 295-339.