World Automation Congress

Third International Symposium on Soft Computing for Industry

Maui, Hawaii
June 11-16, 2000

A Neuro-Fuzzy Approach For Insect Classification

**Habib Gassoumi, Nadipuram R. Prasad,
John J. Ellington and Hung T. Nguyen**

A NEURO-FUZZY APPROACH FOR INSECT CLASSIFICATION

HABIB GASSOUMI and NADIPURAM R. PRASAD
The Klipsch School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, New Mexico USA 88003-8001
hgassoum@nmsu.edu and rprasad@nmsu.edu

JOHN J. ELLINGTON
Department of Entomology, Plant Pathology, and Weed Science
New Mexico State University
Las Cruces, New Mexico USA 88003-8001
joelling@nmsu.edu

HUNG T. NGUYEN
Department of Mathematical Sciences
New Mexico State University
Las Cruces, New Mexico USA 88003-8001
hunguyen@nmsu.edu

ABSTRACT

Control of pesticides in agricultural ecosystems is essential towards minimizing environmental pollution. Lowering the use of pesticides require the implementation of biological control wherein natural enemies and predators are used to control the population of insects that are harmful to the agricultural commodity. In this context, there is a need to develop methods to identify the class of insects that populate specific agricultural ecosystems. This paper discusses a neurofuzzy approach to insect classification. The methodologies discussed in this paper are applied to the classification of insects in cotton farms. A framework for automated Integrated Pest Management Systems is discussed.

KEYWORDS: neurofuzzy systems, soft computing, fuzzy clustering, insect classification, integrated pest management.

INTRODUCTION

A study conducted by the Food and Agricultural Organization (FAO) of the United Nations indicates that the cost of insecticides used worldwide is approximately US $12 billion [1]. Approximately 15% of this amount have been attributed to pesticide use in cotton farming. There are many problems associated with indiscrete pesticide use, most of which have been well documented. These include the resurgence of pest populations after decimation of the natural enemies, development of insecticide-resistant populations, and negative impacts on non-target organisms within and outside the agricultural ecosystem. One of the more serious problems is the development of insecticide resistance. Several insect (pest) species develop resistance to insecticides, and very few chemical control options exist for these pests. Insecticides more adversely affect natural enemies than the target pests. Because predators and parasitoids must search for their prey, they are generally very mobile and spend a considerable amount of time moving across plant tissue. This increases the likelihood that they will be exposed to the insecticide. When insecticides are applied, ideally only the target pests should be affected. The objective therefore is to maximize pest mortality while minimizing harm to natural enemies.

The goal in Integrated Pest Management (IPM) systems therefore, is to implement methodologies that can minimize pesticide use while maximizing the use of biological controls. Achieving this goal requires carefully monitoring the populations of specific insect species and determining the appropriate mix of pesticide and biological controls for the target agricultural ecosystem. In the following sections, we discuss work performed in designing classifier systems that are suitable for IPM systems.

FRAMEWORK FOR IPM SYSTEMS

Statistical approaches to classifying insects have met with limited success. This is due primarily to the fact that spatial patterns of insects are not fixed. This problem is made even more complex by changes in insect population with season and time. Factors such as weather, host plants, soil, predator, parasitoid, and behavioral factors of the insect population can all contribute towards an extremely complex problem in insect classification. Hence it is difficult, if not impossible, to obtain satisfactory statistical models to predict or classify insect categories. In the most recent study [2], a statistical approach was used to classify eight different insect types. While the results are satisfactory in classifying some insect types, results for other types are marginal.

Figure 1 illustrates a framework for developing Integrated Pest Management systems.

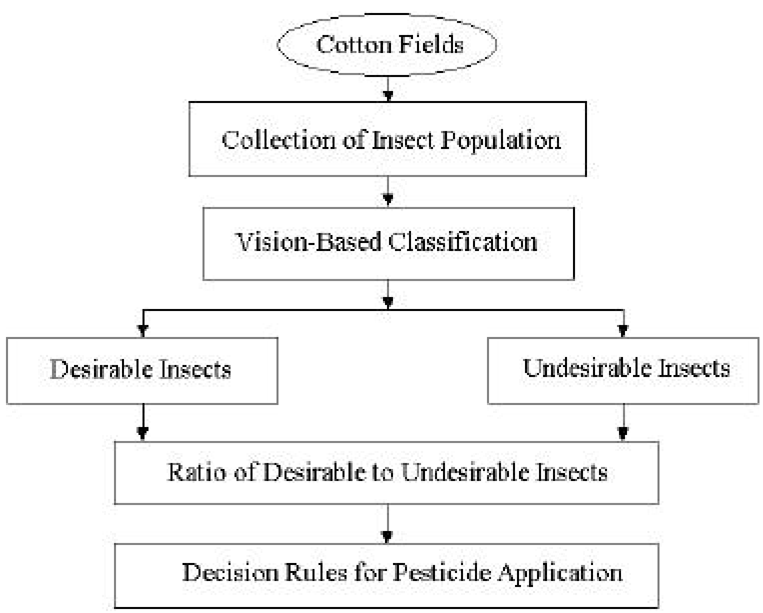


Figure 1. Overview of an IPM system

Referring to Figure 1, the collection of insects is a random sample which is representative of the various insect types that may be present in the agricultural setting such as a large cotton farm. As such, statistical methods for analyzing the data could be applied. However as shown below, the rules for pesticide application is rather linguistic in nature. It is proposed that a computer vision-based approach be used for classifying insects into two basic classes, namely, *desirable* insects and *undesirable* insects. By *desirable,* we mean insects that belong to the predator species, and *undesirable* implying pests. In addition, we are also suggesting that each of the basic classes be further divided into specific insect types that form the biological control species and pest species, respectively. Intuitively then, a ratio of desirable to undesirable insects should provide a basis for developing decision rules for pesticide application. It is clear that these decision rules would be fuzzy *IF-THEN* rules. For example, a typical set of fuzzy rules might be of the form:

> *IF* ***Ratio*** is *Small* and ***Pest x*** is *High THEN* apply *High* concentration ***pesticide Z***
> *IF* ***Ratio*** is *Medium* and ***Pest x*** is *Medium THEN* apply *Medium* concentration ***pesticide Z***

Here *Small*, *Medium*, and *High* are fuzzy subsets that define the variables ***Ratio***, ***Pest x***, and the concentration of ***Pesticide Z*** over their expected range of intervals. Clearly then, some form of defuzzification would result in a decision to apply a certain concentration of ***Pesticide Z***.

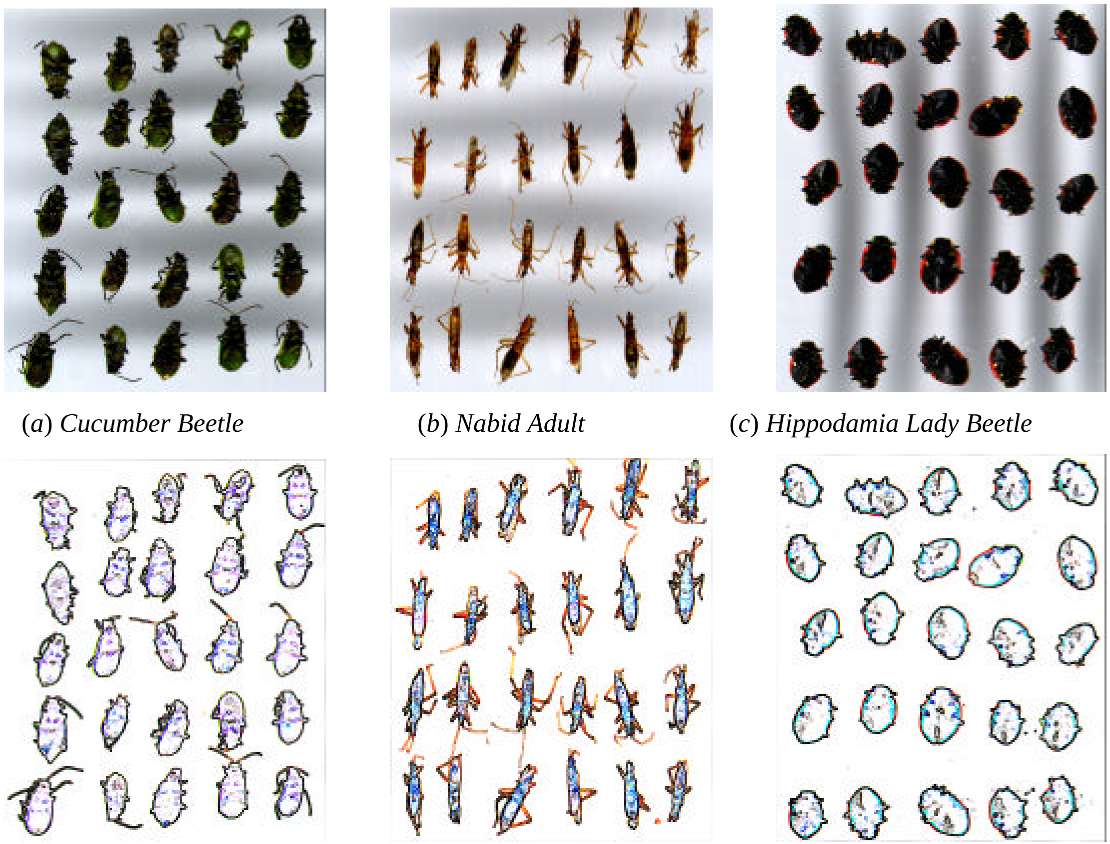
While this paper does not specifically discuss the formulation of such decision-making rules, we are merely suggesting that it is conceivable that such rules can be generated. The significant attributes of such decision rules is based upon the accuracy of insect classification and how well we can actually determine the pest populations in terms of specific species. Note also that the objective is to perform such classification in an automated manner with little or no human intervention. The intent is to develop field deployable IPM systems where there is little or no access to experts in entomology. As such, the IPM system is expected to provide the expertise to the farmer directly in an agricultural setting. Classical statistical methods fail to provide the necessary basis for such an approach to developing successful IPM systems.

Table 1 provides a list of insects that inhabit cotton and alfalfa agricultural ecosystems, and are referred to by their common names.

Table 1. List of Insects in Cotton and Alfalfa Fields

Insect Type	Good	Bad	Comments
Assassin Bug	X		Harmful to humans
Big-eyed Bug	X		Harmless
Green Lace-wing Adult	X		Harmless
Lacewing Larva	X		Harmless
Hippodamia Lady Beetle Adult	X		Harmless
Hippodamia Ladybug Larva	X		Harmless
Nabid Adult	X		Harmless
Stinkbug Adult	X		Harmless
Collops Beetle		X	Harmless
Leaf Hopper		X	Can be destructive
Lygus Adult		X	Destructive
Three-corned Alfalfa Hopper		X	Destructive
Cucumber beetle		X	Destructive

The list includes those which are pests (i.e., have a destructive effect on the crop) and others which are harmless to cotton plants but are considered as biological predators. A few of the insect types listed in Table 1 are shown in Figures 2(*a*)-(*c*).



(*a*) *Cucumber Beetle* (*b*) *Nabid Adult* (*c*) *Hippodamia Lady Beetle*

(*d*) (*e*) (*f*)

Figure 2. Scanned and edge detected images

In Figures 2(*d*)-(*f*), we illustrate the corresponding edge detected images that are suitable for the application of image processing techniques. It is easy to observe that not all insect classes exhibit distinct differences except the Nabid adult shown in Figures 2(*b*) and 2(*e*). The other objects have several similarities in terms of their shapes which can create overlapping features that are difficult, if not impossible, to separate. Note that the “bands” in the images are artifacts created while scanning the insects using a flatbed scanner suggesting the need for a high resolution camera and a good lighting source.

SOFT COMPUTING APPROACH

Soft computing techniques are ideally suited for problems that are far too complex to be defined mathematically. In this context, there is a need to consider a model-free approach to insect classification, whereby the application of neural networks, fuzzy logic, and other hybrid neurofuzzy structures can be justified and used effectively. The primary motivation for adopting this approach is due to the issue of class separability.

While artificial neural networks are universal approximators that are ideally suited for applications in model-free systems, the requirement is that the patterns must be either linearly or nonlinearly separable. Preliminary work using neural networks met with minimal success in that only 2 insect categories could be classified correctly [3]. The small size and similarities in the shape of insects place limitations on the ability of neural networks to produce better results.

Feature Extraction

Table 2 provides a set of typical features computed for 13 insect types.

Table 2. Typical Feature Set for Various Insect Types

Insect Type	Features				
	AR	**P**	**FF**	**RF**	**CF**
Assassin bug	2.4773	109	2.5156	0.2549	0.5049
Big-Eyed bug	2.0556	37	4.9074	0.4972	0.7051
Collops beetle	1.7586	51	5.2493	0.5319	0.7293
Cucumber Beetle	1.7000	85	5.0970	0.5164	0.7186
G L Wing Adult	4.5000	99	1.8056	0.1829	0.4277
H Lady Beetle Adult	1.4375	69	6.4607	0.6546	0.8091
H Ladybug Larva	2.1935	68	3.6090	0.3657	0.6047
Leaf Hopper	3.1071	87	2.7873	0.2824	0.5314
Lacewing Larva	2.5000	60	3.4566	0.3502	0.5918
Lygus Adult	2.4400	61	3.6473	0.3696	0.6079
Nabid Adult	2.8000	70	3.1217	0.3163	0.5624
Stinkbug Adult	1.4028	101	6.7639	0.6853	0.8278
T-C Alfalfa Hopper	1.6136	71	5.0224	0.5089	0.7134

The features shown in Table 2 are computed using the following empirical relationships [4]:

$$\textit{Form Factor}\ (FF) = (4\pi A / P^2) \tag{1}$$

$$\textit{Roundness Factor}\ (RF) = 4A / (\pi D_{\max}^2) \tag{2}$$

$$\textit{Aspect Ratio}\ (AR) = D_{\max} / D_{\min} \tag{3}$$

$$\textit{Compactness Factor}\ (CF) = \sqrt{((4/\pi)A} \ / D_{\max} \tag{4}$$

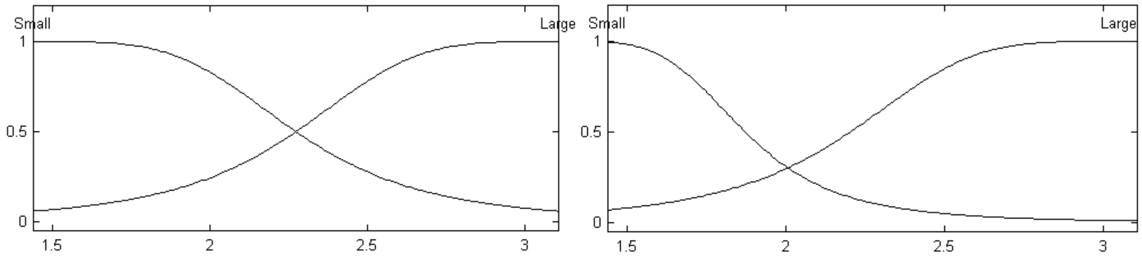
where, A is the Area of object in pixels
P is the Perimeter of object in pixels
$D_{\max}$ and $D_{\min}$ are lengths of the major and minor axes of an ellipse fitted around an object. Note that the Aspect Ratio is 1.0 for a circular object.

Adaptive Neural Fuzzy Inference System (ANFIS) [5]

The structure of ANFIS is very similar to neural networks. In this approach, data clusters are partitioned optimally, and a set of fuzzy *IF-THEN* rules is generated. These rules provide a basis for pattern classification. ANFIS utilizes the least-squares method and the back-propagation gradient descent for identifying linear and nonlinear parameters, respectively, in a Sugeno-type fuzzy inference system. In ANFIS, rules are of the form:

$$Rule^j : IF\ x_1\ is\ A_1^j\ and\ x_2\ is\ A_2^j\ and \cdots and\ x_n\ is\ A_n^j$$
$$THEN\ y = f_j = c_0^j + c_1^j x_1 + c_2^j x_2^2 + \cdots + c_n^j x_n^N$$

where x_i is the input variable, y is the output variable, A_1^j are linguistic terms of the precondition part with membership functions $\mu_{A_i^j}(x_i)$, and where $c_i^j \in R$ are coefficients of linear equations $f_j = (x_1, x_2, \cdots, x_n)$ for $j = 1,2,\cdots,M$ and $i = 1,2,\cdots,n$. Fuzzy subsets in the form of membership functions used by ANFIS are shown below. The features provided in Table 2 are the inputs needed to train the ANFIS. An example of the initial and trained membership functions for one of the input variables is illustrated in Figure 3.



(*i*) *Input Variable AR* (*i*) *Trained Input Variable AR*

Figure 3. Input and trained membership functions in ANFIS

Preliminary results of pattern classification for the insect types shown in Table 1, yield nearly 100% correct classification rates. While, this result is extremely encouraging, training and testing the ANFIS with a large number of samples is required in order to gain the confidence for high success rates in classification. Such high performance in classification will enable the development of field deployable IPM systems.

CONCLUSIONS

In this paper, we have presented a framework for Integrated Pest Management Systems. The need for and the requirements for a successful IPM are highlighted. Some preliminary results using the Adaptive Network Fuzzy Inference System (ANFIS) are presented. These results indicate a strong possibility that highly successful classification systems can be developed. Results of good classification will provide a sound basis for pesticide management and consequently lead to reliable automated Integrated Pest Management Systems.

REFERENCES

1. The Food and Agriculture Organization of the United Nations, (FAO, May 1999) available at: http://www.fao.org/waicent/ois/press_ne/presseng/1999/pren9932.htm
2. Morris, R. L. "Bulk Classification of Insects by Computer Vision." *Master's Thesis*, New Mexico State University, (May 1998).
3. Gassoumi, H., Ellington, J. J., Nguyen, H. T., Prasad, N. R., 1999. A Soft Computing Approach to Insects Classification in the Cotton Field, MIF'99: Proceedings of the International Symposium on Medical Informatics and Fuzzy Technology. Hanoi, Vietnam, 454-485.
4. Russ, J. C., 1999. *The Image Processing Handbook*, CRC Press, Second Edition.
5. Jang, J. S. R., Sun, C. T., Mizutani, E., 1996. "Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence." Prentice Hall (1996).