ADDGEO: An intelligent agent to assist geologist finding petroleum in offshore lands

Ana Cristina Bicharra Garcia
bicharra@dcc.ic.uff.br

Paula Marisa Maciel
paula@dcc.ic.uff.br

Inhaúma Neves Ferraz
ferraz@dcc.ic.uff.br

Instituto de Computação
Universidade Federal Fluminense

Abstract

Part of the work in finding petroleum areas consists of interpreting thin section taken from the sample rocks. Geologists uses many techniques to identifying the rock constituents, but one of the most efficient method relies on recognizing the visual pattern of the thin section using a microscope. This image recognition is a complex task that requires a lot of training.

In this paper, we present an intelligent agent (ADDGEO) that assists geologists identifying the rocks constituents by either guiding the user's analyses or by suggesting interpretation based on the results of a neural net. ADDGEO is an hybrid tool to thin section pattern recognition. It In addition to presenting a model, we describes ADDGEO's system showing the potential of the tool.

Keywords

Knowledge base systems, neural nets, ADD, intelligent agent, image recognition.

Introduction

Most of Brazilian oil reservoir lay on offshore areas. For this reason, a great effort on the Brazilian oil companies have been focused on technologies to make feasible finding and exploring these rich and complex fields.

The study to identify the potential oil places is very expensive per se because it involves getting samples of the undersea rocks, perform sound and nuclear experiments. Although obtaining the data to analyze the field is very expensive, a wrong or even misleading diagnosis may be catastrophic, in economic terms. From the sample rocks, geologists produce thin sections with relevant material to be interpreted using special microscopes. The task is to identify the soil components, oil existence probability and amount, depositional environment, rock permeability and extraction difficulties. In summary, geologists, based on a visual analysis, diagnosis the potential amount of oil and the difficulty to explore it in a specific field. In addition to the visual interpretation, other methods are used to assist the diagnosis such as electric, electromagnetic, seism, spectrometric Alfa to ratify or rectify the visual hypotheses.

The visual thin section analysis consists of a very effective method, but subjective and dependent on the expert's point of view. In addition to being argued by its subjectivity, the thin section analysis method is in danger due to the limited number of experts in the area.

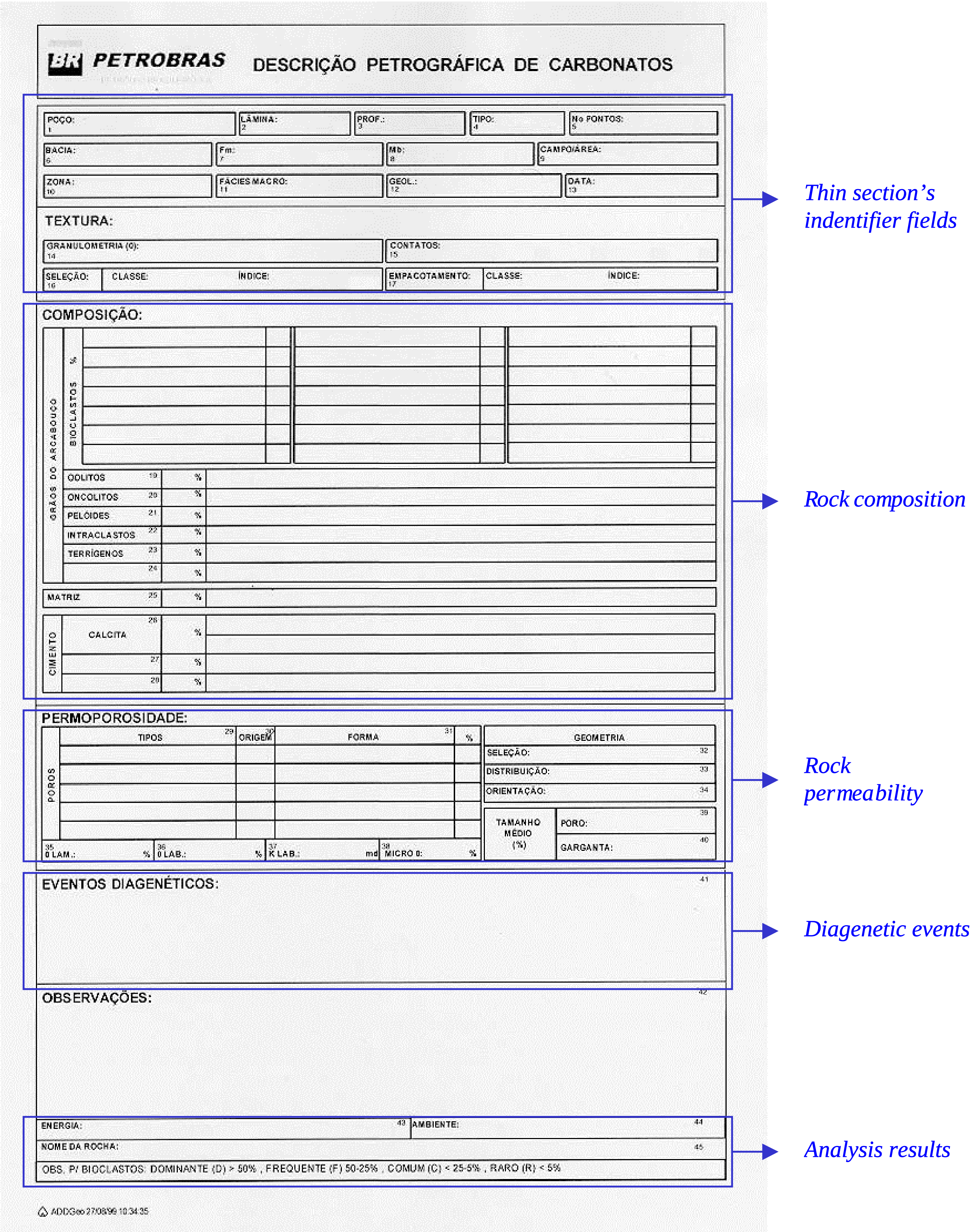
The company's expertise availability in conjunction with the need to maintain this know-how "in house" created a good opportunity to build an intelligent assistant agent based on the knowledge systems and neural networks technologies.

This paper presents a hybrid agent (symbolic and connectionist) based on the ADD model [1]. We also presented a prototype system applied to diagnosis carbonate rocks on Brazilian offshore areas. Initial result have shown very efficient assisting novice to recognize rocks.

Describing the Task

When we study the potential of an oil field, many wheels are made to gather material for analysis. Thin sections are prepared, in the company's laboratory, to contain relevant sample material taken from those exploration wheels. The prepared thin sections are sent to experts to identify the rock. The expert's task consists of identifying:

- Type of grain non bioclasto;
- Existence, percentage and type of grain bioclasto;
- Porosity and permeability percentages;
- Cementation, fragmentation, compactation, neomorphism and other diagnetic events; and
- Depositional environment.



PETROBRAS DESCRIÇÃO PETROGRÁFICA DE CARBONATOS

POÇO: 1 | LÂMINA: 2 | PROF.: 3 | TIPO: 4 | No PONTOS: 5

BACIA: 6 | Fm: 7 | Mb: 8 | CAMPO/ÁREA: 9

ZONA: 10 | FÁCIES MACRO: 11 | GEOL.: 12 | DATA: 13

TEXTURA:

GRANULOMETRIA (0): 14 | CONTATOS: 15

SELEÇÃO: 16 | CLASSE: | ÍNDICE: | EMPACOTAMENTO: 17 | CLASSE: | ÍNDICE:

COMPOSIÇÃO:

GRÃOS DO ARCABOUÇO — BIOCLASTOS %

OOLITOS 19 %

ONCOLITOS 20 %

PELÓIDES 21 %

INTRACLASTOS 22 %

TERRÍGENOS 23 %

24 %

MATRIZ 25 %

CIMENTO — CALCITA 26 %; 27 %; 28 %

PERMOPOROSIDADE:

POROS — TIPOS 29 | ORIGEM 30 | FORMA 31 | %

GEOMETRIA — SELEÇÃO: 32 | DISTRIBUIÇÃO: 33 | ORIENTAÇÃO: 34

TAMANHO MÉDIO (%) — PORO: 39 | GARGANTA: 40

35 Ø LAM.: % | 36 Ø LAB.: % | 37 K LAB.: md | 38 MICRO Ø: %

EVENTOS DIAGENÉTICOS: 41

OBSERVAÇÕES: 42

ENERGIA: 43 | AMBIENTE: 44

NOME DA ROCHA: 45

OBS. P/ BIOCLASTOS: DOMINANTE (D) > 50% , FREQUENTE (F) 50-25% , COMUM (C) < 25-5% , RARO (R) < 5%

ADDGeo 27/08/99 10:34:35

Figure 1: Thin section analysis report standard.

The expert works observing the thin section on a precision microscope and recognizing visual patterns. They generally fill a standard report, as shown in Figure 1, containing the results of their analyses. Generally, their work is done individually or in group.

The ADDGEO Model

After exhaustive interviews with experts in the company, we realized the identification process was subjective and complex, but feasible to be partially automated. A great deal of knowledge could easily become a knowledge base. There are still some parts of the identification were the experts conflict on the term definition (an ontological issue to be investigated in another paper). However, a neural network could be of great help to integrate different perspective identifying bioclastos (life forms).
Figure 2 illustrates the ADDGEO model. There are two ways to interact with the system: automatic diagnosis and cooperative diagnosis,. In the first, the system identifies by itself the rock components through a neural network recognizer. In the other way, users answer questions about what they see in the image and, this way, let the system conclude using its symbolic knowledge model.
The neural net approach is very useful when users have no idea of what is going on. However, the symbolic agent assists users not only to achieve an interpretation, but it guides users learning the domain. In addition of providing and interpretation, the system contains a context sensitive help system that teaches users the meaning of the terms and the logical inferences the system is able to do.

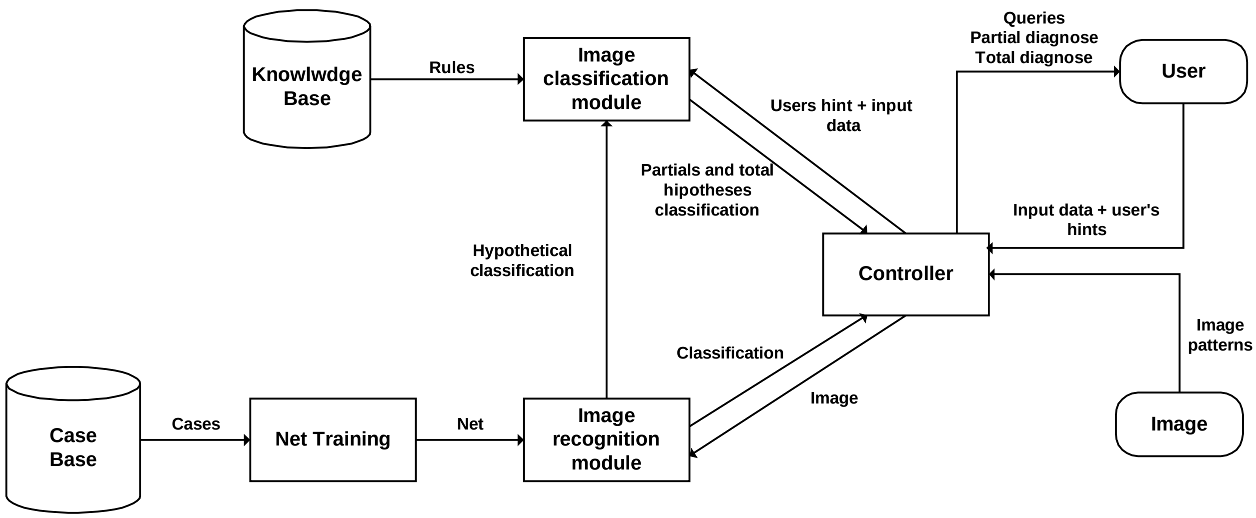


Figure 2: ADDGEO model.

Describing the Knowledge Base

The knowledge in ADDGEO is represented through dependency parametric networks. In this representation, nodes are classification parameters where links represents causal relations. In summary, we used a typical classification model. Figure 3 presents a sample of the representation showing the relation among size, form and attributes to identify possible bioclasto family. Depending on the bioclasto family, other characteristics are considered relevant. Even though, in Figure 3 we show a classification tree with three levels, the entire network contains over 800 parameters.

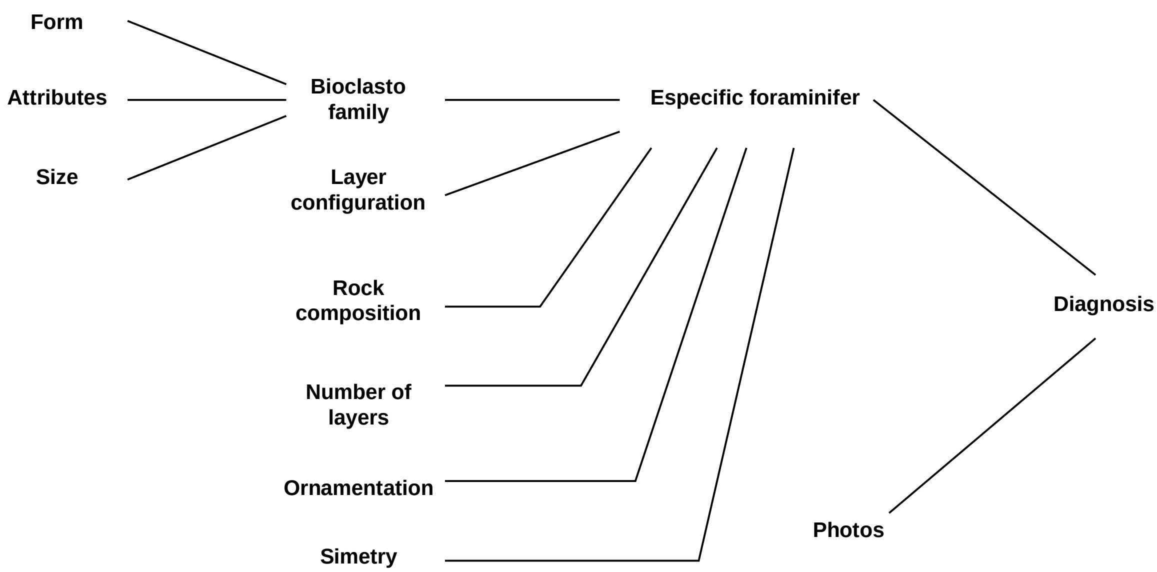


Figure 3: A dependency parametric net sample to identify a specific foraminifer (bioclasto).

Since users are not obliged to provide all input data, the model must consider the existence of unknown values in its inference rules. In the worst case, users provide no information. If the system does not reach a conclusion, it shows the users all open hypotheses with samples of visual data recorded in its database.
The knowledge base grows as users use the system. Users can opt to include a new case in the system's case database. As the case database grows, the system learns a little bit (route learning). From time to time, a data mining or a learning procedure extracts new rules from the recorded cases. Therefore, knowledge don't die.

Describing the Data base

ADDGEO's database are the soul of the system. Even though they are just to support the diagnosis task, they are fundamental to show credibility and to act as the user's extended memory. There are three fundamental databases:

- Image database, containing digitized thin section photos from the all cataloged Brazilian oil fields;
- Grain classification database, containing the taxonomies and meaning of all identifiable grains (bioclasto or not); and
- Thin sections attributes database, containing the attributes that maps a sample to a rock.

We built a friendly query interface to access information in the databases. It is a menu based interface that mounts a query and translates it into a natural language query, so users can rectify or ratify their queries, avoiding the load to learn an SQL commands.

The Neural Net

Neural networks traditionally were applied to classification tasks in simple or linear (one dimension) environments. In 1989 Decatur [2] used neural networks to terrain classification task for terrain radar images. His attempt was a turning point because before this work the preference were directed classical statistical methods (bayesian and nearest-neighbor classifiers). Since 1990 various researchers have presented cases of remotely sensed images recognition. Roli, Serpico and Vernazza [3]described the evolution of these attempts. The back propagation paradigm was the preferred one. In 1993 Hwang et al[4]. Introduced the radial basis function neural network in which the training time is substantially reduced.
The common feature of all these models is the decomposition of the image in a set of different channels, one channel for each sensor and a separate treatment of the channel outputs. Multi layer Perceptrons, according Freeman [5], analyzed each pattern and, in some cases, the output of the channels is combined in a tree-like architecture ending in a majority decoder.

Comparing with the regularity of the plantations, lakes, rivers and roads detected by the remote sensing the "bioclastos" images show the ambition of the ADDGEO Project. This could be done because the neural network is used for hypothesis generation to be confirmed by the built-in expert system of ADDGEO.

In ADDGEO the task was to recognize large bitmaps with ten to fifty thousand 24 bits pixels. And more the images could come from different sources with different resolution and representation (jpeg, bmp, tiff, etc.).

These aspects increase substantially the complexity of the neural network task. In the current version the chosen earning rule and paradigm was the back propagation
The most part of the effort was direct to the image filtering in order to transform bitmaps of thousands of pixels in treatable sets of input data. These goals were achieved using mathematical morphology methods as color segmentation, edge detection and majority code simplification.
The neural network was a three-layer network and we used the Neuralware product [6]. The input layer was fed by the simplification applied to the filtered bitmap.
In the end of simplification task the input data was composed of 625 bits used as input data layer.
The geologists have determined 41 kinds of "bioclatos" to recognize. These classes were codified in the binary method using six bits to the output data layer.

The one and only hidden layer were composed by four elements. The learning cycle involved 100.000 rounds. The output layer content was interpreted as an index for a table entry. This table has the characteristics of the desired "bioclasto".

This task is a seven steps sequence:

1. The user selects a "bioclasto" using the ordinary selection method (fencing the "bioclasto" in a rectangular boundary).
2. The user pushes a button to indicate that he (or she) wants to recognize a pattern.
3. The system applies a sequence of filters to the selected image finding a probable boundary.
4. The probable boundary is simplified and transformed in a bit vectors.
5. The bit vector is used as a neural network input in the pattern recognition task.
6. The neural network processing results in a 6-bit number that is used as a table index.

The table entry obtained is showed to the user identifying the inferred "bioclasto" and exhibiting various images of the same family of the "bioclasto" for reinforce the confidence of the user

The Interaction Interfaces

Based on a study of the task and in the way users realized it, we developed a friendly interface permitting the users execute completely their job. The Figure 4 presents one layout of the main interface showed in the example section. The main menu allows users manipulate their project (thin section to be classified), create alternative diagnoses, consult the databases and obtain help on the system's contents. The interface is divided in four areas:

- Study area, where the user sees the thin section image. Resources like zoom, scale and grid are usable tools in this area;
- Input data area, where the user inputs the thin section identification information;
- Diagnosis area, where the user inputs his classification or asks the system for suggestions or conclusions;
- Working area, where the system interacts with the user in order to obtain a classification. The questions or information presented in this area are classification context dependent.

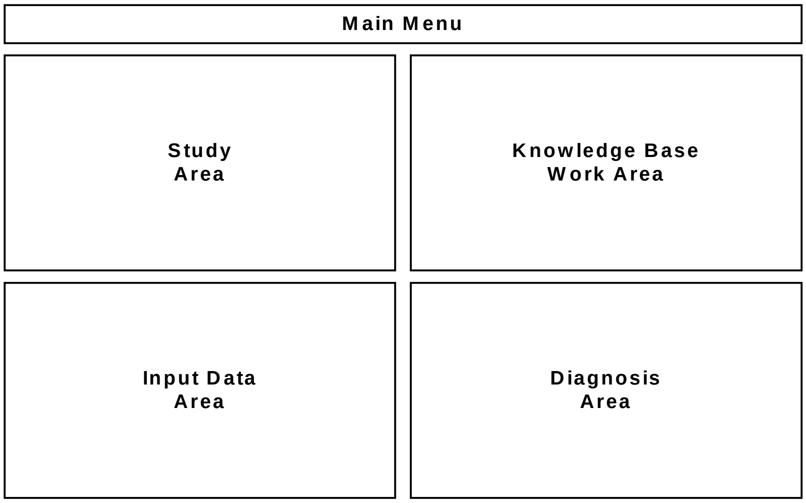


Figure 4: Schematic diagram of ADDGEO system's main interface

Using the ADDGEO system to recognize the constituents in Brazilian carbonate rock's thin cells

Figure 5 presents the main interface of ADDGEO system.

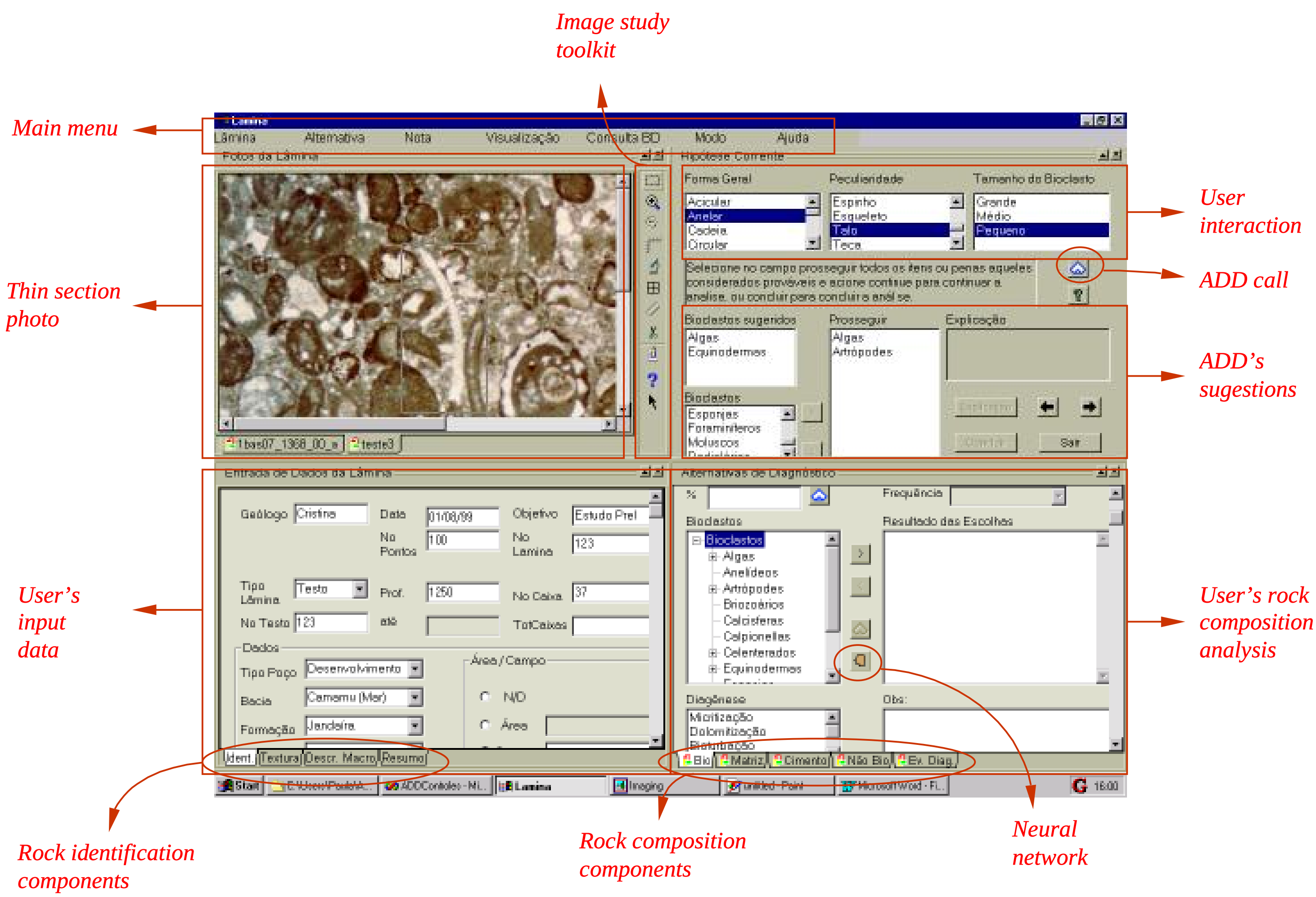


Figure 5: ADDGEO system main interface screen dump

Suppose a geologist receiving a set of thin sections to analyze, interpret and confirm a diagnosis sries about the rock represented by the samples. Using the main interface a new project (or a new case study) is created including all the images of thin section in this case. The images remain available in the left upper part of the interface as showed in Figure 5. The user inputs the thin section general data from the laboratory such as deepness where the sample was extracted, name of the well, well type and the name of the geologist in charge of the analysis. These data are put in the form presented in the left lower part of the main interface.
As soon as the general data have been input the user can begin create alternative diagnosis. Is interesting to note that there are no mandatory fields and the user can omit any field.
The user can create so many study alternatives as he considers necessary and to release any alternative to the Corporate Intranet. Since a diagnosis alternative is created the right upper and right lower parts of the main interface are available to filling.
The right lower part presents the thin section diagnosis. This diagnosis can be made completely by the user or with system help. If the user was able identify the grains he does not need system help. However if the user wants some help he presses the ADD button (presented in Figure 6) and waits for system directives.

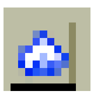


Figure 6: ADDGEO system command button

The system, based in the information furnished by the user, tries to identify what is being asked. If the information are insufficient the system will direct questions to the user until achieve a conclusion. These questions can be in text form or rely in multimedia resources. The interaction between the system and the user takes place in the interface right upper part - the current hypothesis area.
A neural network also can do the help. A button triggers the task of recognition of the image associated to a grain with a head as icon.

Conclusion

We presented a hybrid system (symbolic and neural) to help highly complex images recognition. Our model was applied to off shore carbonated rocks.
Although the system have changed the way the geologists execute this tasks it is been well accepted because it represents a tool for supporting and no to replace those in charge with the tasks.
Now we are collecting data about the rate of help and accuracy of the system in the domain.

References

1. Garcia, A. C. B. – Active Design Documents: A New Approach for Supporting Documentation in Preliminary Routine Design – Ph. D. Thesis – Civil Engineering Department – Stanford University – 1992.
2. Decatur, S. E. – Applications of Neural Networks to Terrain Classifications – Proceedings of the International Joint Conference on Neural Networks 89 – Washington – 1989.
3. Roli, F. , Serpico, S. and Vernazza, G. - – Neural Networks for Classification od Remotely Sensed Images in Fuzzy Logic and Neural Networks Handbook – McGraw Hill Inc. – New York – 1996.
4. Hwang, J. N. , Lay, S. R. and Kiang, R. – Robust Construction Neural Networks for Classification on Remotely Sensed Data - – Proceedings of World Congress on Neural Networks 93 – Portland – 1993.
5. Freeman, J. A. e Skapura, D. M. - Neural Networks Algorithms, Applications and Programming Techniques - Addison-Wesley Publishing Company - Reading - 1991
6. Klimasausker, Casimir C. e Guiner, John P. Neuralworks - Neuralware Inc.- Pitsburgh - 1988.