

# **ADDGEO: um agente inteligente para auxiliar geólogos a interpretar lâminas de rochas carbonáticas**

**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**

## **Resumo**

Reconhecimento de imagens é uma tarefa muito difícil para ser realizada por sistemas computacionais. Muitos esforços têm sido empregado para desenvolver sistemas que realizem esta tarefa. Neste artigo apresentamos um modelo híbrido (simbólico e neural) de reconhecimento de imagem baseada no paradigma de documentação ativa. A complexidade da tarefa é diminuída através da inclusão do usuário como parte do solucionador do problema. Sistema e usuário trabalham de maneira cooperativa para otimizar a realização da tarefa. Foi construído um protótipo de sistema aplicado a interpretação de lâminas de rochas carbonáticas.

## **Palavras-chaves**

Base de conhecimento, rede neural, ADD, reconhecimento de imagem

## **Introdução**

A grande produção de óleo brasileira encontra-se em áreas “offshore”. Há vinte anos o Brasil vem investindo no desenvolvimento de tecnologia de prospecção e exploração em águas profundas. A Bacia de Campos, no litoral do estado do Rio de Janeiro, vem mostrando seu grande potencial petrolífero, mas em águas profundas.

O estudo para se identificarem as bacias potenciais é bastante dispendioso, ainda mais em se tratando de área “offshore”. O número de poços de estudo deve ser reduzido ao mínimo necessário. São feitas lâminas contendo material retirado do testemunho que serão

analisadas por geólogos para identificar o provável ambiente deposicional, tipo de grãos, porosidade e permeabilidade; conseqüentemente, a existência de petróleo e a facilidade de extração do mesmo. Associada à análise da lâmina, outros métodos são utilizados, tais como métodos elétricos, eletromagnéticos, potenciais, perfilagem de poços, sísmica rasa e espectrometria alfa, auxiliando, ratificando e às vezes retificando uma análise.

A análise da lâmina é o teste mais subjetivo, pois é sujeito à identificação do padrão visual da lâmina por especialistas. Além do grau de subjetividade, existe no país um número reduzido de especialistas nessa análise, em especial em casos de lâminas de rochas carbonáticas. No entanto, o conhecimento usado nessas análises é ensinado nas universidades e aprimorado com a experiência na prática. Uma empresa como a Petrobras possui conhecimento acumulado, “know-how”, em vinte anos de trabalho “offshore”. Porém, esse conhecimento ainda está concentrado nos poucos especialistas que vêm se aposentando.

Esse quadro oferece um ambiente propício ao desenvolvimento de bases de conhecimento (ou sistemas especialistas) que organize e torne disponível o conhecimento, não para substituir, mas para auxiliar o trabalho de profissionais da empresa. Esse artigo apresenta o sistema ADDGEO, uma ferramenta de auxílio à interpretação de lâminas de rochas carbonáticas baseado no modelo de documentação ativa ADD [1]. Enfatizamos o modelo do sistema, a tecnologia utilizada e as interfaces de interação com o sistema.

## **Descrevendo a tarefa**

Quando se está estudando a probabilidade de óleo em uma bacia, vários poços de estudos são perfurados e do material coletado são feitas lâminas que serão utilizadas para o estudo. Essas lâminas delgadas, contendo amostras de rochas, são enviadas a especialistas para a identificação dos grãos constituintes e dos eventos diagenéticos que devem ter ocorrido na formação da rocha originária.

A tarefa do especialista consiste em identificar:

- Tipo de grão não bioclasto (como por exemplo: oolito e oncolito)
- Classificação dos bioclastos existentes (como por exemplo: bioclasto-plantônico-globuloso-foraminífero-whitenela)
- Percentual de Porosidade e Permeabilidade
- Presença e ordem dos eventos diagenéticos (por exemplo: cimentação, compactação, dissolução, dolomitização, fraturamento e neomorfismo)
- Identificação do ambiente deposicional (por exemplo: plataforma interna supra-maré)

O especialista trabalha observando as lâminas em um microscópio de precisão e reconhecendo os padrões visuais apresentados. Ele geralmente preenche um relatório padrão, como mostrado na Figura 1, que é o produto de seu trabalho.



## O modelo ADDGEO

Depois de extensas entrevistas com os especialistas na área de identificação de lâminas de rochas carbonáticas, verificamos que o processo, apesar de subjetivo, era factível de gerar um modelo. De posse do modelo da tarefa e do conhecimento do domínio, construímos um agente inteligente que auxilia geólogos, não especialistas.

A Figura 2 ilustra o modelo do sistema. Existem duas formas de interagir com o sistema: diagnóstico automático ou diagnóstico cooperativo. No modo diagnóstico automático, o sistema identifica o objeto através do uso de sua rede neural. No modo diagnóstico cooperativo, o usuário responde a perguntas feitas pelo sistema sobre a imagem da lamina apresentada e, a partir das respostas, o sistema chega a sua conclusão. A rede neural também pode ser usada como ponto de partida da hipótese gerada pelo sistema.

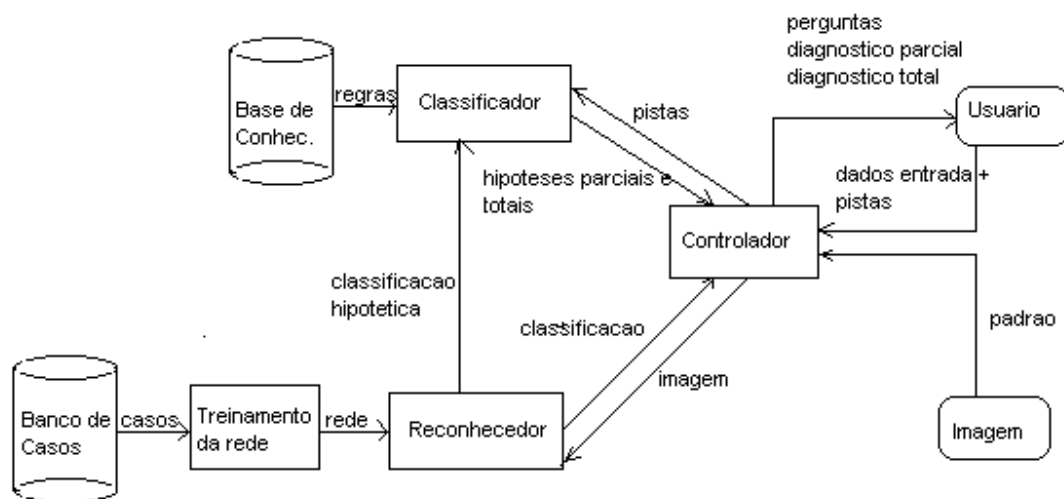


Figura 2: Modelo do sistema ADDGEO

### Base de Conhecimento

A base de conhecimento do sistema ADDGEO está representada por redes paramétricas de associações de características. Por exemplo, a forma, o tamanho e as peculiaridades apresentadas no grão identificam as possíveis classes de bioclastos. Dependendo da classe, um outro conjunto de questões torna-se relevante. Na Figura 3, apresentamos a rede de classificação de bioclastos da família dos foraminíferos. Arranjo das camadas, composição das paredes, numero de camadas, tipo da ornamentação e simetria são características importantes na determinação do tipo específico do foraminífero. Como o usuário não é obrigado a responder a todas as perguntas que o sistema faz, este tem que saber gerar hipóteses com informações incompletas. Caso o sistema não consiga chegar a um diagnostico final, ele apresenta imagens de casos parecidos para que o usuário identifique qual é o mais parecido. Com isso o sistema vai aperfeiçoando seu processo de

diagnostico. Apesar de termos apresentado uma rede com três camadas, o número de camadas para cada classificação pode variar. Atualmente, a rede de classificação do sistema ADDGEO conta com cerca de 800 parâmetros para identificar todos os constituintes da lamina e mais uma sub rede de 50 parâmetros para identificar o provável ambiente deposicional da rocha representada pela lâmina.

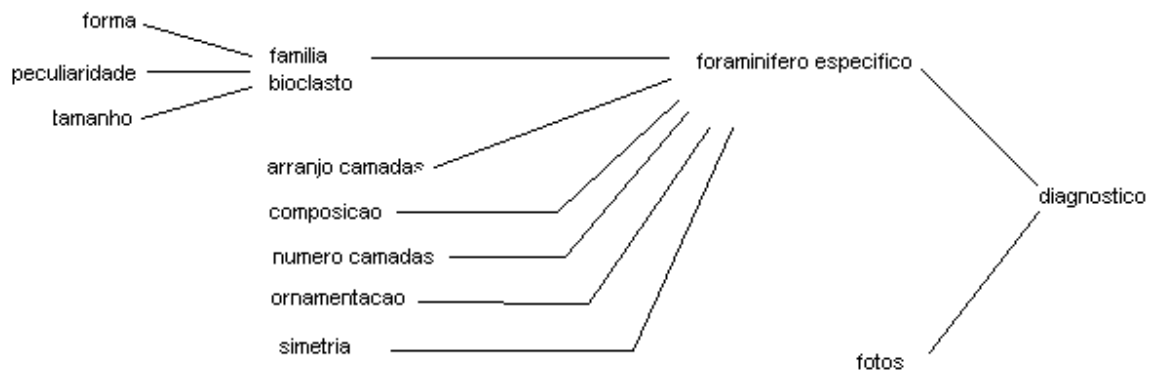


Figure 3: Exemplo da rede paramétrica para reconhecimento de um bioclasto específico— foraminífero.

### **Banco de Dados**

Os bancos de dados no sistema ADDGEO são fundamentais para aumentar a credibilidade do sistema, além de propiciar uma memória estendida ao usuário. Existem três bancos de dados principais:

1. Banco de dados de imagens, contendo um conjunto representativo de fotos de laminas digitalizadas e corretamente classificada. Esse banco de dados é automaticamente alimentado a medida que se usa o sistema;
2. Banco de dados de grãos, contendo a taxinomia dos grãos, bioclastos ou não; e
3. Banco de dados de casos, contendo informações sobre as lâminas, suas características e classificações associadas.

Foi construída uma interface amigável de acesso a esses bancos de dados permitindo que usuários obtenham informações sem precisar saber a sintaxe das linguagens de acesso a banco de dados (SQL).

### **Rede Neural**

As redes neurais tradicionalmente tem sido aplicadas a tarefas de classificação em ambientes simples ou unidimensionais. Em 1989 Decatur [2] empregou redes neurais para a tarefa de classificação de terreno à partir de imagens de radar. Sua tentativa foi um marco pois antes disso a preferencia recaia em métodos estatísticos clássicos (classificadores bayesianos ou de vizinhança mais próxima) . Desde 1990 diversos pesquisadores tem

apresentado casos de reconhecimento de imagens oriundas de sensoriamento remoto. Roli, Serpico e Vernazza [3] descreveram a evolução destas tentativas. O paradigma de retro propagação era o preferido. Em 1993 Hwang [4] e outros introduziram o paradigma de rede neural de função de base radial no qual o tempo de treinamento é substancialmente reduzido.

O ponto em comum de todos esses modelos consiste na decomposição da imagem em um conjunto de distintos canais, cada qual com os dados de um dado sensor e no tratamento separado da saída de cada canal. Os perceptrons em múltiplas camadas, segundo Freeman [5], analisavam cada padrão e, em alguns casos, a saída dos canais era combinada por meio de uma arquitetura em árvore terminando em um decodificador de maioria (o vencedor leva tudo). Considerando que nas imagens oriundas de sensoriamento remoto existe uma regularidade nos padrões a reconhecer (plantações, lagos, rios e estradas, por exemplo) que não ocorre no reconhecimento dos bioclastos pode-se perceber a ambição do Projeto ADDGEO. Isto só pode ser encaminhado porque neste último a rede neural foi empregada para formulação de hipótese a ser confirmada pelo sistema de descoberta de conhecimento intrínseco do Projeto ADDGEO.

No Projeto ADDGEO a tarefa consiste em reconhecer mapas de bits com dezenas de milhares de pixels. Além disso as imagens podem ser oriundas de diversas fontes com distintas resoluções e representações (jpeg, bmp, tiff, etc.). Essas questões aumentam em muito a complexidade da tarefa sendo tratada por uma rede neural.

Na versão corrente estamos utilizando uma rede de retro propagação (back propagation). O maior esforço despendido no projeto foi direcionado para a filtragem de imagens para transformar grandes mapas de bits em conjuntos tratáveis de dados. Estes objetivos foram alcançados utilizando métodos usuais de segmentação de cores, detecção de contornos e simplificação binária.

Utilizou-se uma rede neural com três camadas, utilizando o produto Neuralware [6]. A camada de entrada foi alimentada pelo efeito da simplificação binária aplicada aos mapas de bits filtrados. O resultado da simplificação foi representado por vetores de 625 bits que serviram como memória de curto prazo da camada de entrada. Os geólogos haviam determinado a existência de 41 tipos de bioclastos a reconhecer. Estas classes foram codificadas de maneira binária utilizando seis elementos na camada de saída.

Uma única camada oculta foi otimizada com quatro elementos de processamento. O ciclo de treinamento consistiu de 100.000 iterações. A produção da camada de saída foi interpretado como índice para entrada em uma tabela. Esta tabela contém as características do bioclasto desejado.

Esta tarefa foi realizada em sete passos:

1. usuário seleciona um bioclasto utilizando os métodos usuais de seleção (recorte em vizinhança retangular)
2. usuário aciona um botão indicando a intenção de reconhecer um padrão.
3. sistema aplica uma sequência de filtros à imagem selecionada encontrando um contorno provável.
4. contorno provável é simplificado e transformado em um vetor de bits.
5. vetor de bits é utilizado como entrada de rede neural para a tarefa de reconhecimento.
6. A saída da rede neural é apresentada sob a forma de um número de 6 bits que é usado como índice de uma tabela cuja entrada é constituída de um nome de bioclasto, exibido ao usuário como sendo o bioclasto inferido.
7. sistema mostra ao usuário várias imagens do mesmo tipo de bioclasto para reforço da consistência da inferência.

### Interfaces

A partir do estudo da tarefa e de como os usuários a executavam, desenvolvemos uma interface amigável que permite usuários a realizar inteiramente seu trabalho. Figura 4 apresenta um esquema da interface principal, que será mostrada na seção do exemplo. O menu principal permite que usuários manipulem seu arquivo (lâmina a ser classificada), criem alternativas de diagnóstico, consultem os bancos de dados ou obtenham ajuda quanto ao conteúdo ou uso do sistema. A interface está dividida em quatro áreas:

1. Lâmina: onde o usuário visualiza a imagem da lamina. Recursos como zoom, régua, e grade são ferramentas utilizáveis nessa área;
2. Entrada de Dados: onde o usuário fornece as informações identificadoras da lâmina
3. Diagnóstico, onde o usuário ou entra suas classificações ou pede para que o sistema sugira conclusões;
4. Área de Trabalho: onde o sistema interage com o usuário para concluir sobre uma classificação. As informações ou perguntas apresentadas nessa área são dependentes do contexto da classificação.

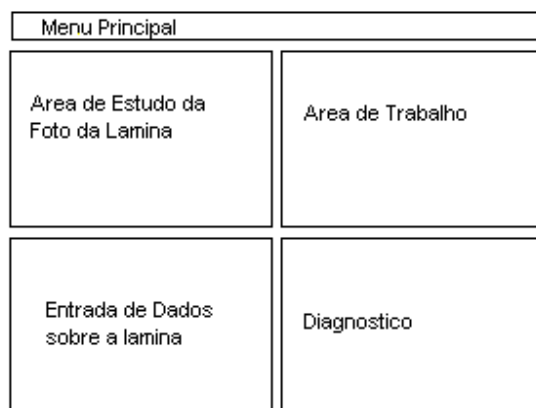


Figura 4: Diagrama esquemático da interface principal do sistema ADDGEO.

## Um exemplo de uso do sistema ADDGEO

Figura 5 apresenta a interface principal do sistema ADDGEO. A seguir descrevemos um cenário de uso do sistema.

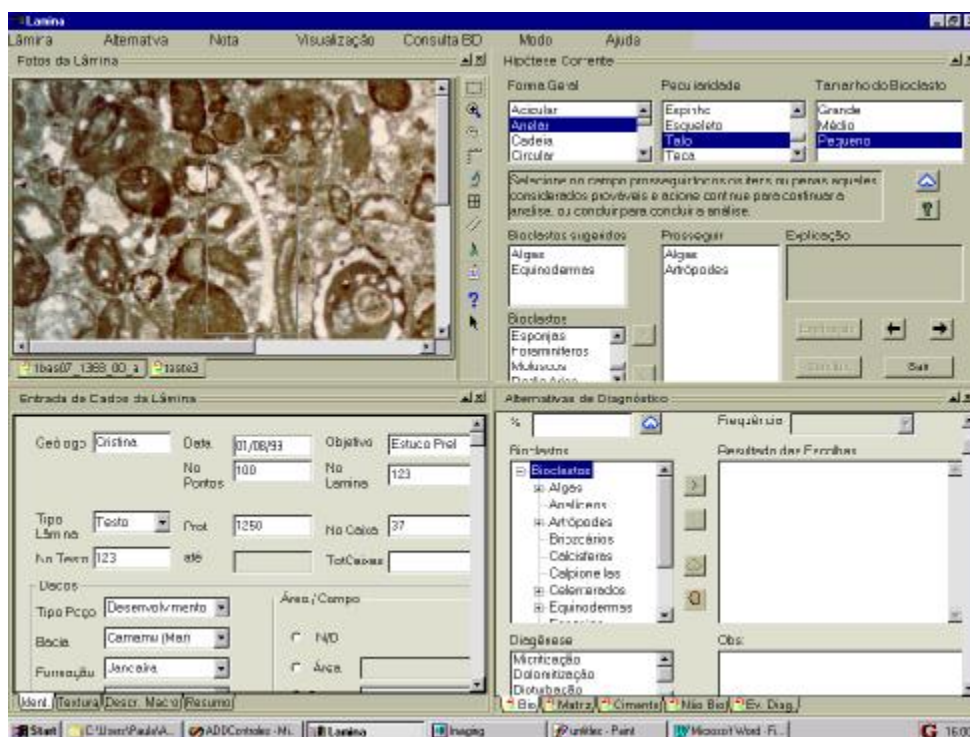


Figura 5: Screen dump da interface principal do sistema ADDGEO.

Imagine que um geólogo recebeu um conjunto de lâminas para analisar, interpretar e conferir uma série de diagnósticos sobre a rocha representada pela amostragem. Através da interface principal, ele criaria um novo projeto (caso) e incluiria todas as imagens das lâminas inseridas no caso. As imagens ficariam disponíveis na parte superior esquerda, como mostrado na Figura 5. O usuário entra com os dados gerais da lâmina vindos do laboratório, tais como: profundidade em que foi retirada a amostra, nome do poço de óleo perfurado e tipo de poço; assim como o nome do geólogo responsável pela análise. Esses dados são inseridos através do formulário apresentado na parte inferior esquerda da interface principal.

Assim que os dados gerais forem inseridos, note-se que não existe obrigatoriedade de preenchimento de qualquer campo, o usuário pode começar a criar alternativas de diagnóstico. Ele pode criar quantas alternativas de estudo ele achar necessárias e disponibilizar qualquer delas na Intranet da empresa. A partir do momento que uma alternativa de diagnóstico é criada, as áreas superior e inferior a direita da interface principal fica disponíveis para preenchimento.



A parte inferior direita apresenta o diagnóstico da lâmina. Tal diagnóstico pode ser preenchido exclusivamente pelo usuário ou auxiliado pelo sistema. Se o usuário conseguiu identificar o grão, ele não precisa de ajuda do sistema. Contudo se ele quer alguma ajuda ele aperta a tecla ADD (representada pela figura 6) e espera por direcionamentos do sistema.



Figura 6: botão de acionamento do sistema ADDGEO

O sistema, baseado na informação disponibilizada pelo usuário, tenta identificar o que está sendo pedido. Caso ele não possua informações suficientes, o sistema fará perguntas ao usuário para chegar a alguma conclusão. Tais perguntas podem ser textuais ou contar com recursos multimídia. Essa parte interativa entre sistema e usuário se dá na parte superior direita—área da Hipótese corrente.

O auxílio também pode ser feito através de uma rede neural. Tal auxílio de reconhecimento da imagem associada a um grão é acionado através do botão com a figura de uma cabeça.

## Conclusões

Apresentamos um sistema híbrido (simbólico e neural) para auxílio a reconhecimento de imagens altamente detalhadas. Nosso modelo foi aplicado para o domínio de reconhecimento de lâminas de rochas carbonáticas em áreas “offshore”.

Apesar do sistema ter mudado a maneira com que os geólogos realizam suas tarefas, ele está sendo bem aceito por representar uma ferramenta de auxílio e não substituição dos encarregados nessa tarefa.

Estamos coletando dados sobre o grau de auxílio e acerto do sistema no domínio.

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 of 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.- Pittsburgh - 1988.