

Reconhecimento de Imagens usando Redes Neurais ou Reflexões sobre a Importância da Filtragem de dados para Reconhecimento de Imagens

Inhaúma Neves Ferraz, Ana Cristina Bicharra Garcia
Instituto de Computação
Universidade Federal Fluminense
Rua. Passo da Pátria, 156 Sala 326 do Bloco E
Niterói – RJ – Brasil
E-mails: ferraz@dcc.ic.uff.br, bicharra@ dcc.ic.uff.br

Abstract

This work presents a comparative study using different neural data models applied to image recognition. We conduct this study using the oil reservoir identification domain. The task consisted of identifying fossils embedded in rock thin sections using a database of scanned images. Our study has shown the impact of image preprocessing (image filtering) in the quality of the recognition task. The use of a sub symbolical model was derived from the user's difficulties identifying the domain taxonomy.

1. Introdução

O presente trabalho foi realizado dentro de um projeto que teve por finalidade a identificação dos componentes de rocha encontrados em uma lâmina delgada recolhida de material extraído de rochas em poços offshore. O projeto foi desenvolvido usando a tecnologia ADD (Garcia, 1992) que permite ao usuário interagir com o sistema na realização de uma tarefa. No caso, a tarefa consiste em identificar nas imagens contidas em uma lâmina delgada:

- Existência, proporção e tipo de grãos bioclastos (forma de vida fossilizada);
- Tipos de grãos não “bioclastos;”
- Porosidade da rocha;
- Eventos diagenéticos;
- Ambiente deposicional.

Construiu-se uma base de conhecimento contendo as heurísticas necessárias para, a partir de uma interação do tipo “Pergunta-Resposta” com o usuário, chegar-se a conclusões sobre os componentes da rocha. Apesar de subjetiva, conseguiu-se um conjunto de heurísticas eficientes para a tarefa, exceto para a identificação de fósseis. Isso motivou o estudo da abordagem conexionista para o reconhecimento dos mesmos. Esse artigo discute o uso de diversos modelos de redes neurais para identificação de fósseis através das imagens

capturadas de lâminas delgadas de rocha. Vale ressaltar que o grande desafio foi a filtragem da imagem.

2. Descrição da tarefa

A tarefa consistiu no reconhecimento de mapas de bits de tamanhos variando entre dez e cinquenta mil bits pixels (com 24 bits cada). As imagens foram obtidas de diversas fontes, sendo que suas resoluções e representações (jpeg, bmp, tiff, etc.) também eram desuniformes. Isto fez com que a maior parte dos esforços fosse direcionada à filtragem de imagens para transformar mapas de bits de milhares de pixels em um conjunto tratável de dados de entrada. Para tanto utilizamos métodos de morfologia matemática tais como segmentação de cores, detecção de contornos e simplificação por decodificador de maioria.

O reconhecimento de imagens por meio de Redes Neurais foi tratado como um processo que compreende os seguintes estágios:

- Obtenção das imagens;
- Tratamento (Filtragem) das imagens;
- Análise do resultado das filtragens;
- Separação das imagens em conjuntos de treinamento e de testes;
- Geração de entrada para a rede;
- Escolha da rede neural mais adequada;
- Treinamento da rede neural;
- Incorporação do código da rede ao aplicativo.

A seguir serão descritos os referidos estágios.

3. Obtenção de imagens

As imagens dos padrões usados para treinar as redes foram originadas através de “scanner” de fotos obtidas de microscópio e fotos de livros, como ilustrada na Figura 1. Não houve uniformidade de escala ou resolução considerando a inexistência de fontes disponíveis para diversas classes de bioclastos que se pretendia reconhecer.

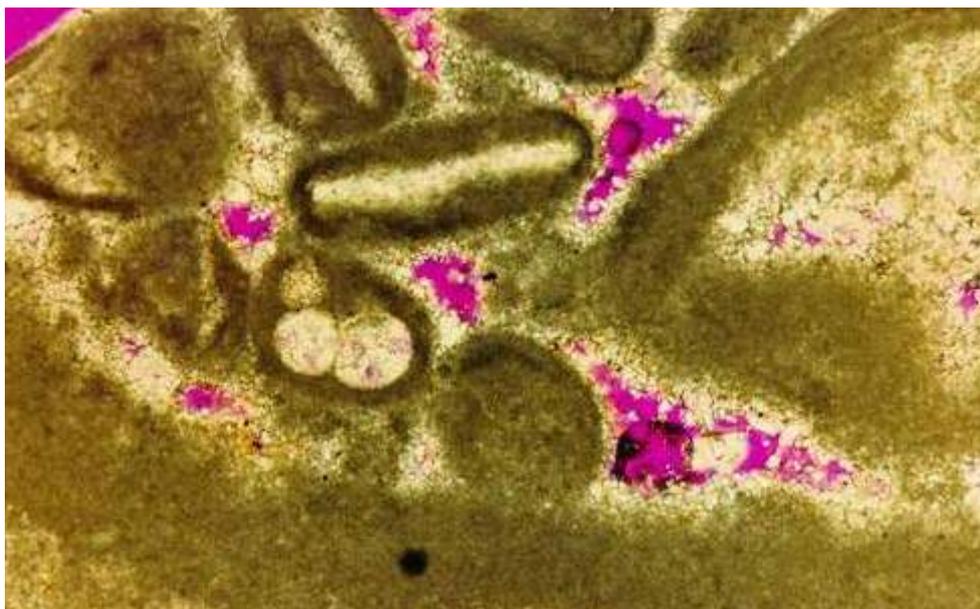


Figura 1: Foto de lâmina contendo um grão bioclasto a ser identificado.

4. Tratamento de imagens

O reconhecimento de padrões visuais deve ser analisado no contexto de Visão Computacional. No atual estado da arte não existe modelo geral de percepção visual aplicável a todas as situações práticas. O trabalho de interpretação de imagens é encarado utilizando-se um conjunto de algoritmos específicos, cada qual realizando uma tarefa limitada. Estes algoritmos envolvem filtros de contraste, detecção de fronteiras de objetos, segmentação da imagem em regiões e outros. Na maioria dos casos usa-se uma seqüência de algoritmos que gera uma ferramenta útil apenas em um conjunto específico de imagens.

As redes neurais tradicionalmente vêm sendo aplicadas em tarefas de classificação em ambientes simples ou lineares (uma dimensão). Em 1989 Dacatur [2] usou redes neurais em tarefas de classificação de terreno através de imagens de radar. Seu trabalho foi um marco significativo, pois antes disso, essa tarefa recebia a preferência de métodos estatísticos clássicos (“bayseanos” e classificadores de vizinhança mais próxima).

Desde 1990 vários pesquisadores têm se dedicado ao reconhecimento de imagens obtidas de sensoriamento remoto. Roli, Serpico e Vernazza [3] descreveram a evolução dessas tentativas. O paradigma de retro propagação tem sido escolhido com mais frequência nestes casos. Em 1993, Hwang[4] apresentou a rede neural de função de base radial na qual o tempo de treinamento é substancialmente reduzido. Contudo, o ponto em comum destas experiências foi a decomposição da imagem em um conjunto de diferentes canais, um para cada sensor com tratamento separado das saídas de cada canal. Segundo Freeman [5], perceptrons multicamadas analisavam cada padrão e, em

alguns casos, a saída do canal era combinada em uma arquitetura em árvore que terminava por decodificadores de maioria. Em comparação com a regularidade das plantações, lagos, rodovias e rios detectados por sensoriamento remoto, as imagens de bioclastos mostram a enorme ambição do projeto em epígrafe.

O diferencial do problema de reconhecimento de lâminas do presente caso constitui na cardinalidade da amostra. O desafio proposto configura-se no reconhecimento de padrões identificados por **cerca de 30.000 bits de entrada**. Os problemas usuais de análise sísmica, reconhecimento de cardiogramas, otimização da época de aquisição de “commodities” tratam de configurações de 40 a 180 bits, no máximo. A utilização Redes Neurais no sensoriamento remoto é feita para amostras de 6 a 12 bits. O desafio do trabalho é, portanto, ampliar a escala da resolução de problemas. Esta ousadia só pode ser tentada porque a rede neural foi utilizada para geração de hipóteses a ser confirmada pelo agente inteligente embutido no sistema.

4.1. Segmentação

O processo de segmentação é um processo de eliminação dos bits de mais baixa ordem, de modo a reduzir a complexidade de representação de imagens. Este processo é fundamental para permitir a detecção de contorno pela transição de cores de pontos. Sem a segmentação, o número de transições é muito elevado e o processo de expansão de regiões de mesma cor interrompe-se com muita frequência não chegando a configurar as figuras desejadas. Usualmente, recebe-se a imagem codificada em RGB. A segmentação pode ser feita pelos seguintes meios:

- Por regiões – Seleciona-se o número de bits a eliminar na simplificação;

- Por cores – Efetua-se uma mudança de coordenadas de RGB para a representação por iluminância, e duas crominâncias (por exemplo, matiz e saturação) e faz-se a segmentação nas novas coordenadas;
- Por intensidade – Efetua-se a mudança de coordenadas para tons de cinza e faz-se a segmentação.

4.2. Detecção de contorno

O algoritmo de detecção de contorno ocorre a partir de um ponto inicial (trabalhando com algoritmos de primeira ordem foi escolhido o algoritmo de Sobel). Quando for escolhida a detecção de contorno genérica, a varredura cobre toda a imagem iniciando no canto superior esquerdo e avançando da esquerda para a direita e de cima para baixo. Sempre que for detectada transição de cor será marcado um ponto de transição (isto se faz quando o usuário não se atreve a distinguir entre pontos interiores e pontos exteriores de um bioclasto).

Filtros locais são aqueles nos quais o ponto inicial é inicializado por um “hot spot” identificado por um clique de mouse em um ponto selecionado pelo usuário.

A varredura ocorre em torno do ponto inicial por meio de expansão da zona selecionada. Sempre que for detectada transição de cor o processo se interromperá no ponto de transição de cor (filtros locais são utilizados quando o usuário pode distinguir, pelo menos, pontos internos de pontos externos).

4.3. Filtros locais

Os filtros locais poderão ser dos seguintes tipos:

- Contorno - Exibe o contorno da região de mesma cor do ponto selecionado sobre fundo branco; e
- Polivalente - Exibe a região da mesma cor do ponto selecionado na cor original sobre fundo branco. Este filtro local faz um pré-processamento não necessitando de segmentação prévia. Na realidade este filtro é uma combinação de:
 - Segmentação por intensidade com número de níveis igual a dois;
 - Contorno;
 - Recorte de imagem ajustando-a ao tamanho da tela.

4.4. Simplificação

O mapa de bits contendo o contorno é dividido em um conjunto de retângulos (foram testados 16 x 16 retângulos e 25 x 25 retângulos). Para cada retângulo, um decodificador de maioria transforma o conjunto de bits do retângulo em um só bit que é transformado em valor de um elemento do vetor de entrada para a rede neural.

5. Análise do resultado das filtragens

A aplicação dos filtros foi feita por tentativas, pois havia imagens que apresentavam melhores resultados quando segmentadas por regiões e outras quando segmentadas por cores.

Foram feitas algumas observações interessantes:

- Dentre as amostras de imagens, havia uma separação de agrupamentos ou “clusters” de imagens para um mesmo bioclasto; e
- Alguns bioclastos apresentavam, em algumas imagens, contornos semelhantes a bioclastos distintos.

Em conseqüência verificou-se a evidência da inadequação da utilização exclusiva do reconhecimento sub simbólico para o problema em questão.

6. Comparação de eficiência

Testamos seis técnicas diferentes de redes neurais ao problema em questão, que são:

- **Retro propagação** - A retro propagação é a mais conhecida técnica de solução do problema de atribuição de créditos proposta por Minsky e Papert.
- **LVQ** - As redes de Aprendizagem por quantização de vetor (LVQ) são redes de classificação de padrões criadas por Teuvo Kohonen que associam vetores a uma de muitas classes e que utilizam aprendizagem competitiva. A aprendizagem é não supervisionada.
- **PNN** - Uma rede neural probabilística (PNN) fornece uma técnica geral para problemas de classificação. Um vetor de entrada, chamado o vetor de características, é usado para determinar uma categoria. As redes PNN usam os dados de treinamento para desenvolver funções de distribuição que, por sua vez, são usadas para estimar a chance de um vetor de características pertencer a uma dada categoria. As PNN são implementação dos classificadores de Bayes. Os estimadores de Parzen constroem as funções de densidade de probabilidade exigidas pela teoria de Bayes. As PNN têm um tempo de aprendizado muito menor do que as redes de retro propagação, mas exigem um grande número de exemplos para o treinamento.
- **Base Radial** - Uma Rede de Função de Base Radial, ou “Radial Basis Function Network (RBFN)” é qualquer rede que possua uma representação interna dos elementos de processamento ocultos (que são as unidades de padrões) com simetria radial.
- **SOM** - Mapeamentos Auto Organizados ou “Self-Organizing Maps (SOM)” dividem itens em categorias apropriadas de objetos similares.

As redes SOM criam um mapeamento bidimensional de características de dados de entrada de uma maneira tal que a ordem é preservada. As redes SOM podem ser usadas como “front end” de redes de retro propagação ou de redes para categorização ou classificação. O treinamento das redes SOM é não supervisionado sendo auto-organizado.

- **MNN** - As redes “Modular Neural Networks” ou “Adaptive Mixtures of Local Experts” consistem em um grupo de redes (ou “local experts”) competindo na aprendizagem dos diversos aspectos de um problema. Uma rede de portagem controla a competição e aprende as diversas regiões como atribuir distintas regiões do espaço de dados a diferentes especialistas locais.

Considerou-se quarenta e uma categorias diferentes de padrões. A codificação de padrões de saída foi binária e, assim, a camada correspondente foi criada com seis elementos de processamento. A camada de entrada, tendo em vista as considerações sobre tratamento da imagem de entrada, foi eleita com seiscentos e vinte e cinco elementos de processamento.

O número de camadas ocultas e o número de elementos de processamento em cada uma delas variou de acordo com os paradigmas escolhidos em cada caso.

7. Resultados obtidos

Foram feitas diversas tentativas e comparações cujos resultados médios podem ser vistos no quadro abaixo que não exhibe as tentativas descartadas logo nos primeiros resultados desfavoráveis (625 x 20 x 6, 625 x 10 x 6, 625 x 12 x 6).

Desta comparação resultou que apenas o paradigma Rede Neural Probabilística (PNN) não convergiu e que os melhores resultados foram obtidos com uma Rede Neural de retro propagação com seiscentos e vinte e cinco elementos na camada de entrada, quatro elementos em uma única camada oculta e seis elementos na camada de saída. A não convergência da Rede PNN não é incomum. Este tipo de rede é muito eficiente em treinamento rápido, mas necessita de grande massa de dados para treinamento, coisa que não ocorreu neste projeto.

Paradigma	Elementos por camada				Treinamento		Erros
	Entrada	Oculta 1	Oculta 2	Saída	Número de Ciclos	Duração	
Retro prop.	625	50	-	6	200.000	2hs: 09 m	46
Retro prop.	625	4	-	6	50.000	0:04 m	45
Retro prop.	625	4	-	6	Auto treino	0:03 m	124
Retro prop.	625	2	-	6	50.000	0:04 m	58
Retro prop.	625	50	12	6	200.000	0:03 m	47
LVQ	625	50	-	6	11.000	2hs: 31 m	56
PNN	625	54	-	6	7.000	0:04 m	Inconclusivo
Base Radial	625	54	-	6	20.000	0:02 m	56
SOM	625	10	10	6	7.000	0:01 m	56
MNN	625	1	-	6	50.000	0:11 m	46

8. Conclusões

A motivação de usarmos processamento sub simbólico para atacar o problema de reconhecimento de bioclastos nas imagens de lâminas delgadas veio pela dificuldade em identificar a semântica da terminologia entrada de dados para o processamento simbólico. Por não existir um padrão de terminologia nas características descritoras das imagens, usuários não familiarizados com o jargão não sentem confortáveis com as analogias identificadoras de forma ou padrão.

Usamos vários modelos sub simbólicos para automatizar a tarefa de identificação das imagens. Constatamos que, indiferente do tipo de modelo usado, a grande dificuldade do reconhecimento dos padrões não estava no processamento em si, mas no pré-processamento exigido; isto é na filtragem da imagem.

Observando-se os resultados obtidos verificamos que o índice de reconhecimento foi extremamente baixo, índice de acerto em torno de 30%, o que não é aceitável.

Estes resultados são opostos aos resultados obtidos no reconhecimento de padrões geológicos com leis de variação lineares representadas por linhas (reconhecimento de litofácies). Nestes casos a mesma equipe, com o mesmo “software”, conseguiu índices de acerto sempre superiores a 80% (oitenta por cento) nas mesmas bacias geológicas. Aparentemente a escala do problema extrapolou as condições de ambiente propícias a seu emprego.

Para melhorar os resultados obtidos duas possíveis linhas de ação:

- Adotar um enfoque probabilístico pelo qual deve-se escolher de maneira pseudo-aleatória um elevado número de pontos de início de busca de contorno e

buscar reconhecimento por decodificador de maioria; ou

- Melhorar o processo de filtragem das imagens atuando sobre as etapas de segmentação e detecção de contorno.

A primeira linha de ação foi tentada no prosseguimento do projeto de agente inteligente e seus resultados não propiciaram uma melhora significativa.

Tendo em vista que o grande problema está sendo o de limpar as imagens, estamos estudando duas possibilidades: melhorar os filtros das imagens computacionalmente e colocar o usuário como parte do modelo no processo de filtragem do dado. Talvez seja mais fácil para ele perceber as impurezas da imagem que estão tornando o reconhecimento automático impossível.

Referências

- [1] Garcia, A. C. B., Ferraz, I. N. e Maciel, P. M. - ADDGEO: An intelligent agent to assist geologist finding petroleum in offshore lands - Proceedings of V Intelligent Systems Applications - New Orleans - 2000.
- [2] Decatur, S. E. - Applications of Neural Networks to Terrain Classifications - Proceedings of the International Joint Conference on Neural Networks 89 - Washington - 1989.
- [3] Decatur, S. E. - Applications of Neural Networks to Terrain Classifications - Proceedings of the International Joint Conference on Neural Networks 89 - Washington - 1989.
- [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