

Título do Projeto:

Reconhecimento de Minerais Translúcidos de Rochas Ígneas no Microscópio Petrográfico Utilizando Inteligência Artificial.

Grupo desenvolvedor:

Matheus Henrique Schaly.

Tema de pesquisa:

O problema principal que a projeto pretende minimizar é o custo envolvido na identificação de minerais translúcidos de rochas ígneas visto em microscópio petrográfico. O custo compreende a aquisição de mais quantidades de microscópios, mais tempo necessário para a identificação mineral, e, por fim, mais tempo despendido para obter conhecimento necessário para o reconhecimento do mineral.

Objetivo de pesquisa:

O objetivo da pesquisa é automatizar e agilizar a identificação de minerais translúcidos de rochas ígneas através do microscópio petrográfico.

Justificativa:

Caso bem-sucedido, a identificação de minerais será feita com mais rapidez e com melhor precisão.

Além disso, a necessidade da aprendizagem de identificação de tais minerais poderá ser reduzida, já que o software fará o trabalho, restando mais tempo para pesquisas em outras áreas, nas quais máquinas são menos capazes de realizar.

Ademais, microscópios petrográficos serão menos necessários, logo, a quantidade de microscópios poderá ser reduzida, assim como o espaço que eles ocupam, ocasionando menos gasto de capital na compra de microscópios e na sua infraestrutura.

Outra vantagem apresentada, sendo que o programa seja capaz de realizar sua tarefa com boa precisão, a utilização da divisão laboratorial referente ao estudo petrográfico de lâminas delgadas, não terá mais tanta significância, quando tratar-se de identificação mineral.¹ Já que, após a finalização do software, qualquer pessoa que possua um microscópio petrográfico, e consiga retirar dele algumas imagens (que serão de fácil captura), será capaz de reconhecer um mineral translúcido sem a necessidade do envio para um laboratório (ou ainda grande conhecimento sobre a área), com isso, contribuindo com a rapidez do processo e a economia de recursos financeiros.

Ainda, acredita-se que o programa possa ser capaz de identificar, um mineral translúcido, com melhor precisão do que um ser humano. Geólogos com vasta experiência em mineralogia, conseguem identificar a maioria dos minerais (minerais mais raros apresentam maior dificuldade, já que são menos vistos), mas as vezes encontram-se em discussão com seus colegas referente a classificação de um mineral específico. E isso não acontece pôr divergências de categorizações de mineral, já que os minerais apresentam características definidas em livros. Sendo assim, a causa da discussão pode ser originada pelo fato de cada um ter um ponto de vista diferente. Por exemplo, algum deles pode pensar que a cor que está sendo vista é um azul-claro, enquanto o outro classificaria (a

mesma imagem) como um cinza-claro. Ou ainda, questões sobre relevo do mineral, birrefringência, pleocroísmo... podem ser fonte de divergência. Essas discordâncias podem ser amenizadas com o uso do programa.

Finalmente, com o aperfeiçoamento do programa, talvez haja a possibilidade de estender a análise não somente para minerais translúcidos, mas também avançar para opacos, assim como feições geológicas encontradas nas seções delgadas, como o ambiente geológico, sua evolução, seu tipo de desenvolvimento com o passar do tempo, entre outras coisas. Que só podem ser analisadas em amostras de mão (que possuem certas limitações, como, por exemplo, não é possível identificar minerais microscópios em amostras de mão) e de lâminas delgadas.

O estudo de rochas e minerais com a ajuda do microscópio petrográfico é um procedimento obrigatório para quase todos os tipos de trabalho que o geólogo desenvolve. Com este instrumento é possível observar aspectos que, devido à sua reduzida dimensão, não podem ser observados nas amostras de mão. Por exemplo, detectam-se minerais de pequenas dimensões, observam-se os contatos entre diferentes minerais e pode estimar-se a sua percentagem numa dada rocha.²

Exemplo de rocha ígnea em microscópio petrográfico:

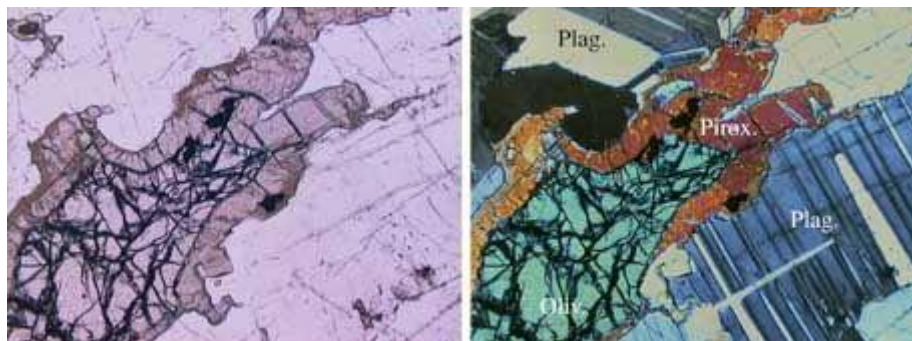


Figura 1: Microfotografias em nicóis paralelos (luz natural) e nicóis cruzados (polarizador inserido). Minerais da rocha: Plag. = Plagioclásio, Pirox. = Piroxênio, Oliv. = Olivina.³

Trabalhos correlatos:

“Deep Learning With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.”⁴ Tal revisão, que será descrita abaixo, foi feita em 2013.

Programas de aprendizagem profunda tentam imitar a atividade em camadas dos neurônios no neocórtex, que é 80% do cérebro onde o pensamento ocorre. O programa aprende, de um modo muito real, a reconhecer padrões em representações digitais de sons, imagens e outros dados.

A ideia de que o programa pode simular a grande variedade de neurônios do neocórtex em uma “rede neural” artificial tem décadas de idade, e tem levado a decepções em relação a sua inovação. Mas por conta das melhorias nas fórmulas matemáticas e nos computadores cada vez mais poderosos, os cientistas da computação podem agora modelar muitas mais camadas de neurônios virtuais do que nunca.

Com essa maior profundidade, eles estão produzindo avanços notáveis no reconhecimento de fala e imagem. No último mês de junho, um sistema de aprendizado profundo que foi mostrado 10 milhões de imagens dos vídeos do YouTube provou-se duas vezes melhor do que qualquer esforço anterior de reconhecimento de imagem para identificar objetos como gatos.

O Google em particular, tornou-se um imã para aprendizagem profunda e talento relacionado a AI. Em março, a empresa comprou uma startup co-fundada por Geoffrey Hinton, professor de ciência da computação da Universidade de Toronto, [...] diz que planeja “tirar ideias deste campo e aplica-las a problemas reais”, como reconhecimento de imagens, pesquisa e compreensão da linguagem natural, diz ele.

Em relação a redes neurais: Um programa mapeia um conjunto de neurônios virtuais e então atribui valores numéricos aleatórios, ou “pesos”, às conexões entre eles. Esses pesos determinam como cada neurônio simulado responde (com uma saída matemática entre 0 e 1) a uma característica digitalizada, como uma borda ou uma tonalidade de azul em uma imagem, ou um nível de energia específico em uma frequência em um fonema, a unidade individual do som em sílabas faladas.

Os programadores treinariam uma rede neural para detectar um objeto ou um fonema atacando a rede com versões digitalizadas de imagens contendo tais objetos, ou ondas de som contendo tais fonemas. Se a rede não reconhecer com precisão um determinado padrão, um algoritmo ajustaria os pesos. O objetivo final deste treinamento era conseguir que a rede reconhecesse consistentemente os padrões na fala ou conjuntos de imagens que nós, humanos, conhecemos como, digamos, o fonema “d” ou a imagem de um cão. Isto é muito parecido com a maneira que uma criança aprende o que é um cachorro, notando os detalhes da forma da cabeça, do comportamento, e do gosto em animais peludos, animais que latem que outras pessoas chamam de cachorro.

Em 2006, Hinton desenvolveu uma maneira mais eficiente de ensinar camadas individuais de neurônios. A primeira camada aprende características primitivas, como uma borda em uma imagem ou a mais ínfima unidade de som de fala. Ele faz isso encontrando combinações de pixels digitalizados ou ondas sonoras que ocorrem com mais frequência do que deveriam ocorrer por acaso. Uma vez que a camada reconhece exatamente esses recursos, eles são alimentados para a próxima camada, que se treina para reconhecer recursos mais complexos, como um canto ou uma combinação de sons de fala. O processo é repetido em camadas sucessivas até que o sistema possa reconhecer fielmente fonemas ou objetos.

Como gatos. Em junho do ano passado (junho de 2012), o Google demonstrou uma das maiores redes neurais já vistas, com mais de um bilhão de conexões. Uma equipe liderada pelo professor de ciência da computação de Stanford, Andrew Ng, e pelo pesquisador do Google, Jeff Dean, mostrou as imagens ao 10 milhões de imagens aleatórias de vídeos do Youtube. Um neurônio no modelo de programa fixou em imagens de gato. Outros focaram em rostos humanos, flores amarelas, e outros objetos. E graças ao poder do aprendizado profundo, o sistema identificou esses objetos discretos mesmo que nenhum humano os tivesse definido ou rotulado.

O que surpreendeu alguns especialistas em IA, porém, foi a magnitude da melhora no reconhecimento de imagens. O sistema categorizou corretamente objetos e temas nas imagens do YouTube 16% do tempo. Isso pode não parecer impressionante, mas foi 70% melhor do que os métodos anteriores. [...] havia 22.000 categorias para escolher; Colocando corretamente os objetos em algumas das categorias, sendo que algumas das categorias, por exemplo, era necessário distinguir entre duas variedades semelhantes de arraia. Isso teria sido um desafio mesmo para a maioria dos humanos. Quando o sistema foi solicitado a classificar as imagens em 1.000 categorias mais gerais, a taxa de precisão saltou acima de 50 por cento.

O treinamento de tantas camadas de neurônios virtuais no experimento usou 16.000 processadores de computador, o tipo de infraestrutura de computação que o Google desenvolveu para seu mecanismo de busca e outros serviços. Pelo menos 80% dos recentes avanços em AI podem ser atribuídos à disponibilidade de mais poder de computador [...].

No entanto, há mais do que o tamanho dos centros de dados do Google. A aprendizagem profunda também se beneficiou do método da empresa de dividir as tarefas de computação entre muitas máquinas para que elas possam ser feitas muito mais rapidamente [...]. Isso também acelera a formação de redes neurais de aprendizagem profunda, permitindo que o Google execute redes maiores e alimente muito mais dados para as redes neurais.

Os recursos computacionais do Google [...] são cruciais, dizem os defensores da aprendizagem profunda, porque o próprio cérebro é ainda muito mais complexo do que qualquer uma das redes neurais de hoje. “É necessário muitos recursos computacionais para fazer com que as ideias funcionem”, diz Hinton.

Inside DeepMind's latest attempts to achieve a general artificial intelligence: What are progressive neural nets?⁵

Tradicionalmente, uma rede de aprendizagem profunda será treinada através de um aprendizado de reforço profundo (DeepRL, “deep reinforcement learning”) sendo alimentada com enormes quantidades de dados e fornecido o tempo para aprender como executar uma tarefa, como reconhecer os elementos de uma imagem, dominar Space Invaders (jogo da Atari) ou vencer Lee Sedol (campeão mundial do jogo milenar de tabuleiro Go) em Go.

“Cada uma dessas redes neurais é realmente poderosa, e cada uma delas pode atingir um nível sobre-humano nessas tarefas”, diz Hadsell (pesquisador científico), “mas cada rede é separada. Ainda não há uma rede neural que possa ser treinada, simultaneamente, para identificar imagens, jogar Space Invaders e ouvir música.

Google DeepMind pairs with NHS to use machine learning to fight blindness⁶

A Google DeepMind anunciou sua segunda colaboração com a NHS, trabalhando com Moorfield Eye Hospital no leste de Londres, para construir um sistema de aprendizado de máquina que eventualmente será capaz de reconhecer condições ameaçadoras relacionadas a visão a partir, apenas, de uma varredura digital do olho.

[...] Treinar uma rede neural para fazer a avaliação de exames oculares poderia aumentar muito a velocidade, assim como a precisão do diagnóstico, salvando a visão de milhares de pessoas.

[...] “Eu estava lendo sobre o aprendizado profundo e o sucesso que a tecnologia teve no reconhecimento de imagens”, disse Pearse Keane (consultor oftalmologista) quando encontrou um artigo sobre o treinamento da DeepMind para jogar os jogos da Atari, o primeiro sucesso público da empresa.

[...] disse Pearse Keane “eu tive uma ideia brilhante que a aprendizagem profunda poderia ser realmente boa em olhar as imagens de olho. A Tomografia de Coerência Óptica é a minha área, e temos o maior depósito de imagens OCT do mundo [...].

O professor Peng Tee Khaw, chefe do centro de pesquisa oftalmológica de Moorfields, disse que a chave para a colaboração foi o enorme aumento no volume de exames de retina incrivelmente precisos disponíveis. “Esses exames são incrivelmente detalhados, mais detalhados do que qualquer outro exame do corpo que fazemos: podemos ver a nível celular. Mas o problema para nós é lidar com essa quantidade de dados.”⁷

Solução proposta e diferencial:

Primeiramente, deve-se analisar mais o problema, ler artigos geológicos como estes: <http://link.springer.com/article/10.1007/BF02087099> e/ou <https://pdfs.semanticscholar.org/5130/8518bca6f28c3a7cf25116d743d94e5d09b4.pdf>, para tentar achar outras possíveis soluções menos trabalhosas, ou mais úteis para automatizar o processo de identificação mineral. Assim como explorar mais artigos relacionando redes neurais, aprendizagem de máquina, aprendizagem profunda, inteligência artificial e reconhecimento de imagem, a fim de saber se é realmente possível e tangível a solução do problema proposto através desses artifícios relacionados a inteligência artificial.

Após a ideia não ser refutada pela pesquisa na internet e livros, ainda é necessário conversar com profissionais da área de ciência da computação (e similares) e geólogos, com o propósito de ouvir as opiniões deles sobre o projeto, para diminuir os riscos de insucesso, ou ter possíveis melhores soluções. E o diálogo não pode limitar-se para o Brasil, mas com a opinião estrangeira também.

Após esse longo processo, caso a ideia ainda seja válida, a intenção é usar a aprendizagem profunda para ensinar a máquina a reconhecer minerais translúcidos. Com a pouca informação que consegui pesquisar, aparentemente, para o meu objetivo, seria necessário um grande volume de dados, dados estes que devem ser padronizados e serem bem analisados antes de serem submetidos como dados úteis.

Para isso propõem-se às universidades que colaborem a recolher tais dados (novamente, não somente uma universidade, e não somente brasileiras) já que o sucesso de tal programa ajudaria no campo de trabalho deles e não seria algo tão trabalhoso de ser arranjado, pois o uso de microscópio petrográfico é praticamente diário e estudantes o usam com frequência. Sugere-se a utilização de 6 imagens (todas utilizando microscópio petrográfico) para a análise eficiente dos minerais: a primeira seria em luz natural; a segunda, seria também em luz natural, mas girando a platina rotativa à 45°; a terceira

seria voltar 45° e ativar o polarizador (i.e., cruzar os nicóis); a quarta seria girar novamente a 45° (ainda com o polarizador); a quinta seria, na posição atual, ativar o analisador; e a sexta seria apenas desativar o compensador, e ativar a lente de Bertrand.

As propriedades dos minerais em lâmina delgada incluem: cor, pleocroísmo, hábito, clivagem, relevo e alteração, todas essas propriedades são expostas na luz natural. Já com os nicóis cruzados, as propriedades analisadas são: isotropismo, ângulo de extinção, geminação, zoneamento, dispersão, birrefringência e cor de interferência, e figura de interferência.⁸

A junção da primeira com a segunda imagem reconheceria a cor, pleocroísmo, hábito, clivagem e relevo, analisando todas as propriedades da luz natural. Já com a ativação do polarizador, a junção da terceira e quarta imagem identificariam o isotropismo, ângulo de extinção, geminação, zoneamento e dispersão. A quinta imagem reconheceria birrefringência e cor de interferência. E a sexta imagem detectaria a figura de interferência.

Posteriormente, já com essas imagens, a parte mais difícil seria desenvolver o software que as reconhecesse e, posteriormente, classificasse corretamente as imagens fornecidas ao programa. É algo que, ao que tudo indica, é novo, difícil e ainda pouco explorado. As empresas que o desenvolvem, possuem tecnologia de ponta assim como computadores de ponta. Isso seria algo a ser desenvolvido em conjunto com profissionais da área da ciência da computação, incluindo profissionais mais focados em IA (mais especificamente redes neurais, aprendizagem de máquina, aprendizagem profunda e reconhecimento de imagem) e levaria muito tempo para o seu desenvolvimento, levando possivelmente anos para concretizar-se. Portanto, esse seria o ponto do projeto com o maior custo envolvido.

Fundamentação teórica:

Redes Neurais e Aprendizagem Profunda:

Ambos os conceitos estão inseridos dentro de aprendizagem de máquina.

Rede neural é um paradigma da programação inspirado biologicamente, que permite a um computador aprender com dados observacionais.

Aprendizagem profunda é um poderoso conjunto de técnicas de aprendizagem em redes neurais.⁸

A aprendizagem de máquina tradicional baseia-se em redes rasas, composta por uma entrada e uma camada de saída e, no máximo, uma camada oculta entre elas. Entretanto, mais de três camadas (incluindo entrada e saída) qualifica uma rede neural como aprendizagem “profunda”. Portanto, profundo é um termo técnico estritamente definido que significa mais de uma camada oculta.⁹

Redes neurais são programas de computador montados de centenas, milhares ou milhões de células cerebrais artificiais que aprendem e se comportam de maneira notavelmente semelhante aos cérebros humanos. Algo interessante sobre uma rede neural é de que você não precisa programa-la para ela aprender, ela aprende por si mesmo, como um cérebro.

Dentro de um computador, o equivalente a uma célula cerebral é um interruptor chamado transistor.

Os dois principais motivos para o cérebro e o computador pensarem de maneira diferente é, a diferença entre a quantidade de ligações e a quantidade em si, de neurônios e transistores. O cérebro possui 10 a 100 vezes mais celular cerebrais do que a quantidade de transistores em um computador. E, enquanto um transistor tem até duas portas lógicas, neurônios naturais chegam a ligar-se a até 10.000 vizinhos.

A rede neural aprende comparando a saída que uma rede produz com a saída que se pretendia produzir e utiliza a diferença entre eles para modificar os pesos das ligações entre as unidades na rede. Com o tempo, a retropropagação faz com que a rede aprenda, reduzindo a diferença entre a saída real e a saída pretendida até o ponto em que os dois coincidem exatamente, então a rede calcula as coisas exatamente como deveria.¹⁰

Reconhecimento de Imagem:

Visão computacional é um campo interdisciplinar que lida com a forma como os computadores podem ser feitos para ganhar alto nível de compreensão a partir de imagens digitais ou vídeos. Do ponto de vista da engenharia, ele procura automatizar tarefas que o sistema visual humano pode fazer.

As tarefas da visão computacional incluem métodos para aquisição, processamento, análise e compreensão de imagens digitais e, em geral, lidar com a extração de dados de alta dimensão do mundo real para produzir informações numéricas ou simbólicas, por exemplo, nas formas de decisões. Entender neste contexto significa a transformação de imagens visuais (a entrada da retina) em descrições do mundo que pode interagir com outros processos de pensamento e desencadear uma ação apropriada. Esta compreensão da imagem pode ser vista como o desprendimento de informações simbólicas a partir de dados de imagens usando modelos construídos com a ajuda da geometria, física, estatística e teoria da aprendizagem.

Como uma disciplina científica, visão computacional preocupa-se com a teoria por trás de sistemas artificiais que extraem informações de imagens. Os dados de imagem podem assumir várias formas, como sequências de vídeo, visualizações de várias câmeras ou dados multidimensionais de um scanner médico. Como disciplina tecnológica, a visão computacional busca aplicar suas teorias e modelos para a construção de sistemas de visão computacional.

Subdomínios da visão computacional incluem a reconstrução de cena, detecção de eventos, rastreamento de vídeo, reconhecimento de objeto, estimativa de pose de objeto, aprendizagem, indexação, estimativa de movimento e restauração de imagem.¹¹

Neste caso, o subdomínio mais essencial seria o de reconhecimento de objetos. Que é definido por uma tecnologia no campo da visão computacional para encontrar e identificar objetos em uma imagem ou sequência de vídeo. Os seres humanos reconhecem uma infinidade de objetos em imagens com pouco esforço, apesar do fato de que a imagem dos objetos pode variar um pouco em diferentes pontos de vista, em muitos tamanhos e escalas diferentes ou mesmo quando eles são traduzidos ou girados. Os objetos podem até ser reconhecidos quando parcialmente obstruídos da vista. Essa tarefa ainda é um desafio para os sistemas de visão computacional. Muitas abordagens para a tarefa foram implementadas ao longo de várias décadas. Havendo vários métodos característicos para o reconhecimento de imagem.¹²

Classificação da pesquisa:

Estilo “Apresentação de Algo Diferente”, algo antigo feito de uma nova maneira.

O produto é caracterizado ainda em áreas emergentes da computação que é a IA junto ao reconhecimento de imagens.

A pesquisa está apresentando uma forma diferente de resolver um problema, visto que diferentes métodos de reconhecer minerais em microscópio petrográfico não incluem análise de imagens fornecido por um software.

Trabalhos são uma simples comparação entre técnicas, sendo elas, a análise feita em laboratório usando a investigação visual do humano, referente a do software. Porém, exigiria rigor científico na apresentação dos resultados, envolvendo a descrição do processo de desenvolvimento do programa. Comparando a precisão oferecida pelo programa em relação ao ser humano, e isso deveria ser meticulosamente analisado, já que a precisão é um dos pontos cruciais do projeto.

Comparações seriam tanto qualitativas quanto quantitativas, pelo fato da quantidade de dados usados, por exemplo, 10 mil imagens, e qualitativa, apresentando a precisão do ser humano, exemplo 90%, e do programa, exemplo 95%.

Típico em áreas novas (IA junto à reconhecimento de imagem), no qual não se dispõe de grandes bases de dados para testar teorias empiricamente, já que no começo da pesquisa os dados deveriam começar a serem coletados.

Referência bibliográficas:

¹<http://www.corelab.com/ps/petrology>

²<http://www.dct.uminho.pt/rpmic/mic.html>

³<http://www.dct.uminho.pt/mictic/magma.html>

⁴<https://www.technologyreview.com/s/513696/deep-learning>

⁵<http://www.techworld.com/big-data/inside-deepminds-latest-attempts-achieve-general-artificial-intelligence-3646730>

⁶<https://www.theguardian.com/technology/2016/jul/05/google-deepmind-nhs-machine-learning-blindness>

⁷https://profiles.uonbi.ac.ke/cnyamai/files/lecture_4__microscopic_study_of_minerals_under_the_microscope.pdf

⁸<http://neuralnetworksanddeeplearning.com>

⁹<https://deeplearning4j.org/neuralnet-overview#concept>

¹⁰<http://pages.cs.wisc.edu/~bolo/shipyard/neural/local.html>

¹¹http://en.wikipedia.org/wiki/Computer_vision

¹²https://en.wikipedia.org/wiki/Outline_of_object_recognition