

# Identificação de minerais por meio de Redes Neurais Convolucionais: um estudo comparativo entre Inteligência Artificial e o Sistema Visual Humano

Victor Félix Arinos<sup>1</sup>, Thiago Meirelles Ventura<sup>1</sup>, Natali Félix Arinos<sup>2</sup>,  
Amarildo Salina Ruiz<sup>3</sup>

<sup>1</sup>Instituto da Computação – Universidade Federal de Mato Grosso (UFMT)  
Av. Fernando Corrêa da Costa, nº 2367 - Boa Esperança. Cuiabá - MT - 78060-900

<sup>2</sup>Instituto de Física – Universidade Federal de Mato Grosso (UFMT)  
Av. Fernando Corrêa da Costa, nº 2367 - Boa Esperança. Cuiabá - MT - 78060-900

<sup>3</sup>Faculdade de Geociência – Universidade Federal de Mato Grosso (UFMT)  
Av. Fernando Corrêa da Costa, nº 2367 - Boa Esperança. Cuiabá - MT - 78060-900

{victorfelixer, asruiz}@gmail.com, thiago@ic.ufmt.br

talifelix@fisica.ufmt.br

**Abstract.** *Mineral identification is extremely necessary for modern geology, especially when dealing with specific methods of geochronology and extraction of geological information. This process can be time consuming due to human evaluation in identification. In this way, this paper proposes the creation of a Convolutional Neural Network using microscopic images to identify Apatites, Quartz and Zircons, thus reducing manual resources and errors in this process. Identification tests with volunteers were carried out, comparing the results obtained by computer. In the end, the capacity of the Neural Networks of identifying minerals is verified, achieving results superior to human efforts and accuracy above 90%.*

**Resumo.** *A identificação de minerais é extremamente necessária para a geologia moderna, em especial quando lidamos com métodos específicos de geocronologia e extração de informações. Este processo pode consumir um grande período de tempo devido a avaliação humana na identificação. Assim, Este artigo propõe a criação de uma Rede Neural Convolutional utilizando imagens microscópicas para identificação de Apatitas, Quartzos e Zircões, diminuindo os recursos manuais e erros nesta etapa. Além disso, testes de identificação com voluntários foram realizados, comparando os resultados obtidos com computador. Ao final, comprova-se a capacidade das Redes Neurais Convolucionais, com resultados superiores aos esforços humanos e acurácia acima de 90%.*

## 1. Introdução

A classificação de rochas e minerais é um processo trabalhoso da geologia moderna necessário para diversas sub-áreas, como a Engenharia de Minas, Petrologia e outras áreas menores da Geociência [Miynarczuk et al. 2013]. Minerais em geral, como a Apatita ( $Ca_5(PO_4)_3(OH, F, Cl)$ ) e o Zircão ( $ZrSiO_4$ ), costumam ser identificados através

da manipulação humana com lupa, microscopia ou utilizando técnicas mais sofisticadas e custosas, como a difração de Raio-X e Espectroscopia Raman [Izadi et al. 2017, Bayka and Yılmaz 2010].

Estes minerais, em especial, possuem grande relevância para a geocronologia, como no Método de Datação por Traços de Fissão (MTF) e no Método de Datação por U/Pb, ambos métodos de datação mineral. Esta concentração de estudos em torno do MTF aplicado a Apatita, por exemplo, se justifica devido a similaridade em que ocorre o traço de fissão no mineral e o surgimento de petróleo, sendo que ambos os casos ocorrem na mesma temperatura (entre, aproximadamente 50 e 130°C) [Franco 2006]. Além disso, minerais como o Zircão podem ser utilizados no processo de diminuição de lixo nuclear [Zhang 2017].

Contudo, as abordagens de identificação de minerais possuem desvantagens como a subjetividade humana na manipulação para identificação, que torna o processo menos confiável, além do alto custo para aquisição ou utilização da difração de Raio-X e Espectroscopia Raman. Uma das formas de automatizar este processo e aumentar a confiabilidade da análise humana é por meio do desenvolvimento e aplicação de Inteligência Artificial (IA) com o uso de reconhecimento de padrões. Assim, este artigo propõe a criação de uma Rede Neural para identificação de Minerais utilizando Redes Neurais Convolucionais (CNN), além da comparação dos resultados obtidos pela IA e pelos esforços manuais do Sistema Visual Humano.

O artigo está estruturado em cinco seções, sendo a segunda seção dedicada ao referencial teórico e trabalhos relacionados, que abrange questões quanto à Redes Neurais Convolucionais, Data Augmentation e Testes com Humanos; a terceira, discorrendo sobre a metodologia aplicada, sendo os resultados obtidos apresentados na quarta seção. Por último, tem-se as considerações finais e os encaminhamentos para trabalho futuros.

## **2. Referencial Teórico e Trabalhos Relacionados**

Para a realização deste estudo foram aplicadas técnicas de Redes Neurais e *Data augmentation*, sendo que atualmente a utilização de métodos computacionais para a identificação de minerais vem sendo amplamente aplicado como forma de melhorar a precisão de reconhecimento dos grãos analisados. Em especial, tem-se as técnicas de Inteligência Artificial, como a de Redes Neurais Convolucionais [Maitre et al. 2019], caracterizada pela capacidade de resolver problemas utilizando abordagens como o Reconhecimento de Padrões, Clusterização/Categorização, Predição, entre outros [Jain et al. 1996, Borji and Itti 2014] de diversos parâmetros diferentes [Thompson et al. 2001] e objetos [Pereira and Porto 2019].

Quanto a utilização de RNA para identificação de minerais, há trabalhos na área. Thompson, Fueten e Bockus (2001) utilizaram 27 parâmetros diferentes, como textura, saturação e intensidade, para classificar minerais em rochas obtendo resultados de aproximadamente 90% de acurácia. De forma semelhante, Nurdan e Yılmaz (2010) e Izadi, Sadri e Mahdokht utilizaram Redes Neurais para identificar minerais em rochas através da utilização das cores presentes nas imagens, obtendo um resultado superior a 80% e 90% de acurácia, respectivamente. Outros estudos, como o proposto por Maitre, Bouchard e Bedard (2019) utilizaram imagens obtidas diretamente de microscópio para identificação de minerais por meio de diferentes métodos, obtendo resultados acima de

90% de acurácia.

Contudo, percebe-se que nenhum dos trabalhos atuais utiliza o Sistema Visual Humano como parâmetro de comparação, sendo que o processo manual de identificação por esforço humano continua sendo o mais utilizado. Além disso, o único trabalho que utiliza imagens microscópicas não identifica quais minerais especificamente estão sendo analisados, mas somente se estão presentes na imagem ou pertencem a um determinado grupo ou família em comum. Autores como Borji e Itti (2014) utilizaram o reconhecimento de objetos como matéria-prima para a comparação. Seus resultados demonstram que os modelos de IA treinados superavam a capacidade humano de reconhecimento, alcançando resultados superiores com menor esforço. De forma semelhante, Fleuret et al. (2011) estuda e compara a performance de humanos e computadores na classificação de imagens criadas sinteticamente. Neste caso, os esforços humanos alcançaram melhores resultados que a IA, demonstrando uma necessidade de adaptação ou melhoria das tecnologias utilizadas.

### **3. Metodologia**

#### **3.1. Seleção de minerais e obtenção de imagens**

Três minerais foram selecionados para realização do treinamento da Rede Neural, sendo eles: Apatita, Zircão e Quartzo. Esta escolha foi feita devido a importância da Apatita e do Zircão para o Método de Traços de Fissão [Bruand et al. 2014], já que ambos são considerados minerais acessórios, ou seja, se formam a partir de um mineral essencial como o Quartzo devido a variações de pressão e temperatura [Klein and Dutrow 2009]. Além disso, possuem alta capacidade de retenção de informação da sua época de criação, fator importante para o estudo de geocronologia e obtenção de informações geológicas.

A coleta das imagens ocorreu em laboratório, onde todas as amostras já haviam sido previamente classificadas por técnico especialista. Todo o processo foi acompanhado por uma pessoa com experiência na manipulação de minerais através de microscópio, buscando adquirir imagens iguais ou próximas das utilizadas para identificação dos minerais por olho humano. Assim, todas as amostras foram obtidas através do microscópio petrográfico L3030 com objetiva, gerando imagens com os seguintes parâmetros: imagem com resolução 2048x1536 e formato Joint Photographics Experts Group (JPEG); zoom óptico de 20x; máxima iluminação transmitida.

Estes parâmetros foram escolhidos buscando padronizar as imagens, além de simular a identificação por olho humano que, apesar de poder ser feita com outras especificações como zoom diferente, possui melhores resultados nestas condições. Ao final da coleta das imagens obteve-se um total de 383 imagens, sendo 101 de Apatita, 106 de Quartzo e 176 de Zircão.

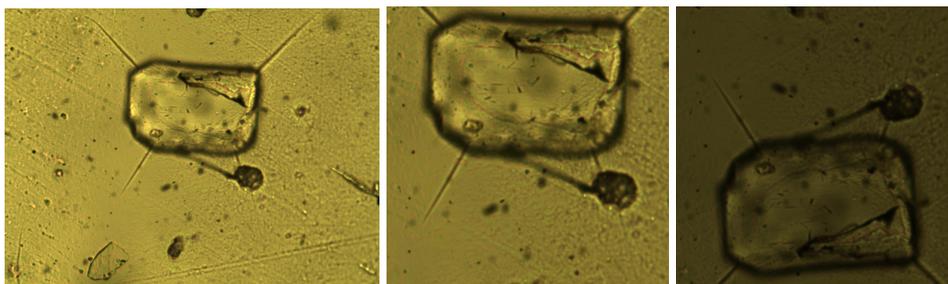
#### **3.2. Data Augmentation e tratamento da base de imagens geradas**

Para aumentar a diversidade de imagens obtidas e remover a discrepância do número de figuras de cada classe, utilizou-se a técnica de *Data Augmentation*. Assim, um determinado grupo de imagens já existente é utilizado para criar novas figuras [Krizhevsky et al. 2012] sem perda de representatividade. Para isso, alguns parâmetros para alteração das imagens foram definidos: rotação para esquerda ou para direita; distorção de largura ou comprimento; aumento ou diminuição de brilho, sendo a variação mínima 0,3 e a máxima 0,9;

cortes em relação ao centro da imagem para a resolução de 1052x1052; inversão da imagem em relação a sua posição. Assim, uma imagem poderia ser alterada em diversos aspectos ao mesmo tempo, sendo que isto ocorreria de acordo com a porcentagem de chance estipulada para cada um dos parâmetros.

Estes parâmetros foram escolhidos buscando simular situações comuns durante o processo de identificação, como mudanças de iluminação, zoom, posição da resina no microscópio e profundidade do mineral na resina. Dois grupos de imagens foram separados: um denominado Teste com 30 imagens de cada classe, totalizando 23% do total de figuras; e outro chamado Treino com 293 imagens, representando 77%.

A etapa de *Data Augmentation* foi aplicada somente no conjunto Treino, mantendo os dados de teste somente com as imagens originais. Cada classe foi manipulada até atingir um total de 320 imagens, totalizando 960 para treino. Conforme mencionado, estas novas imagens possuem variação em sua rotação, distorção, brilho, zoom e posição. Na Figura 1 tem-se na primeira imagem algo obtido através do microscópio e, após aplicação de *Data Augmentation*, variações da mesma imagem. As variações poderiam ser obtidas através do uso de microscópio variando a intensidade da luz, posição da resina e do zoom aplicado, não comprometendo assim a qualidade do estudo.



**Figura 1. Imagem original de uma Apatita e suas respectivas imagens geradas por meio de Data Augmentation**

Ao final, as imagens coletada foram redimensionadas, utilizando a interpolação bicúbica devido a seu baixo custo computacional e grande capacidade de produzir imagens lisas, para três resoluções: 128x128, 64x64 e 32x32. Buscou-se assim analisar a melhor especificação possível para a Rede Neural com o menor custo computacional, além de auxiliar na remoção de características desnecessárias das imagens.

### **3.3. Rede Neural Convolutacional**

A Rede Neural Convolutacional proposta foi criada utilizando o framework Keras, uma biblioteca para implementação de algoritmos de Redes Neurais e *Deep Learning*, sendo desenvolvida na linguagem python. Esta API utiliza como *backend engine* o TensorFlow, podendo ser executada com outras ferramentas [Tokui et al. 2015, Krizhevsky et al. 2012]. Depois de vários testes, a arquitetura selecionada para este trabalho possui duas camadas iniciais de convolução, seguido de três camadas densas com 128, 64 e 3 neurônios. Além disso, as duas primeiras camadas densas foram formadas com a função de ativação ReLU, caracterizada por ser de eficiente e fácil para computação [Krizhevsky et al. 2012], e a última camada densa com a função de ativação softmax para realizar a classificação.

As duas camadas de convolução foram configuradas para aplicar 64 filtros diferentes para cada imagem, sendo que o kernel utilizado possui tamanho 3. Estas especificações foram escolhidas com base em testes prévios onde a adição ou remoção de uma camada de convolução resultou em decréscimo na capacidade de reconhecimento da Rede Neural com base na acurácia. Além disso, tem-se uma camada de agrupamento (MaxPooling) após cada convolução e 1 camada de concatenação (Flatten) antes das camadas densas.

Após a criação da RNA ocorreu o treinamento dos modelos. Cada imagem contida no conjunto de treinamento foram analisadas em 10 épocas diferentes. Como foram geradas três resoluções diferentes de imagens, houve também três treinamentos. Em seguida, os três modelos gerados previamente foram testados com as imagens originais, ou seja, as 90 imagens obtidas diretamente pelo microscópio. Ao final do processo de teste foram recolhidos os parâmetros de acurácia no reconhecimento dos minerais dos três modelos.

### 3.4. Testes de identificação de minerais por meio do Sistema Visual Humano

Após a etapa de criação e treinamento do Modelo de RNA para identificação de minerais, houve o início da etapa de testes com o Sistema Visual Humano. Esta etapa ocorreu como forma de obter dados para comparação da capacidade humana de identificação de minerais com o melhor Modelo gerado.

O teste ocorreu em ambiente controlado e os voluntários foram instruídos inicialmente com uma imagem de cada tipo de mineral a ser identificado como forma de exemplificar o processo. O teste consistia em identificar as 90 imagens originais obtidas através do microscópio sem redimensionamento. Os voluntários, durante todo o processo, puderam consultar as imagens de exemplo. Além disso, o tempo foi cronometrado para mensurar o tempo necessário para identificação.

Foi utilizado um formulário para coletar as respostas. Este formulário possui um campo com a numeração da imagem sendo visualizada e três opções onde o usuário deve assinalar se o mineral atual é uma Apatita, um Quartzo ou um Zircão. No total, nove pessoas participaram do teste. A Tabela 1 estão os dados dos voluntários, no qual a primeira coluna representa sua identificação, seguido da sua escolaridade e formação e, por último, sua experiência com identificação de minerais. Ao final dos testes houve a contabilização do número de erros e acertos de cada voluntário, além do cálculo da porcentagem de acertos.

**Tabela 1. Caracterização dos voluntários**

Ident.	Escolaridade	Formação	Experiência
0	Graduação em andamento	Geologia	Iniciação Científica e disciplinas
1	Graduação em andamento	Física	Iniciação Científica
2	Doutorado em andamento	Física	Pesquisa na pós-graduação
3	Graduação em andamento	Geologia	Iniciação Científica e disciplinas
4	Graduação em andamento	Geologia	Disciplinas
5	Graduação em andamento	Geologia	Disciplinas
6	Graduação em andamento	Geologia	Iniciação Científica e disciplinas
7	Graduação em andamento	Geologia	Iniciação Científica e disciplinas
8	Doutorado em Geologia	Geologia	Docência em Geologia

## 4. Resultados

### 4.1. Modelos gerados e acurácia

Com a etapa de treinamento e teste concluída, com cada uma das três resoluções geradas, houve a análise dos dados de acurácia obtido. Na Tabela 2 estão elencados os resultados obtidos por cada modelo. Na primeira coluna está a identificação do Modelo, em seguida a acurácia durante o treino e, por último, a acurácia no teste. Observa-se que a acurácia de 98.31% foi obtida por dois modelos com resoluções de imagens diferentes. Isto pode ocorrer devido a diminuição de detalhes presentes nas imagens, facilitando a identificação das características principais. Por último, tem-se o Modelo 128x128, cuja acurácia alcançou aproximadamente 92%.

**Tabela 2. Acurácia de cada Modelo**

Modelo	Acurácia - Treino	Acurácia - Teste
32x32	97.81%	98.31%
64x64	99.48 %	98.31%
128x128	98.75%	91.53%

### 4.2. Resultados obtidos pelos voluntários

Após o processo de teste com os voluntários, ocorreu a verificação das respostas fornecidas por meio de gabarito previamente preenchido. Na Tabela 3 estão elencados os dados obtidos pelos testes. Na primeira coluna está a identificação dos voluntários, seguido da quantidade de minerais identificados corretamente (com o máximo de 90) e da porcentagem de acertos. Por último, tem-se o tempo gasto para finalizar o teste.

**Tabela 3. Resultados obtidos do teste de identificação de minerais por humanos no qual tiveram que classificar 90 imagens.**

Voluntário	Quantidade de acertos	Porcentagem de acertos	Tempo (minutos)
0	72	80%	9:41
1	78	86.6%	21:47
2	75	83.3%	7:25
3	64	71.1%	5:33
4	83	92.2%	10:21
5	63	70%	8:07
6	72	80%	12:15
7	76	84.4%	11:17
8	75	83.3%	20:42

Analisando os dados obtidos, observa-se que a média da porcentagem de acertos foi de 83%, com desvio padrão igual a 7%. Além disso, a maior quantidade de acertos foi de 83 (92,2%) amostras identificadas, enquanto a menor foi de 63 (70%). Quanto ao tempo, tem-se que a média dos voluntários foi de aproximadamente 11 minutos, sendo que o maior tempo gasto foi de 21 minutos e 47 segundos e o menor foi de 5 minutos e 33 segundos.

Percebe-se que o maior número de acertos foi realizado por um voluntário com graduação em andamento em Geologia. Sua experiência com identificação de minerais é proveniente das suas disciplinas, em especial de Mineralogia.

### 4.3. Comparações de resultados - Inteligência Artificial e Sistema Visual Humano

Na Tabela 4 estão os dados dos resultados organizados de forma comparativa. Quanto ao número de acertos, pode-se comparar a acurácia da Rede Neural Convolutiva com a porcentagem da média de acertos dos voluntários. Tem-se então que, independente do modelo, a Rede Neural Convolutiva obteve um resultado superior aos esforços humanos. Mesmo no pior caso, com o Modelo 128x128 e sua acurácia de 91,53%, há ainda uma diferença de aproximadamente 10% em relação a média de acertos humanos.

**Tabela 4. Comparação de resultados**

Item avaliativo	Humano	Rede Neural
Melhor porcentagem de acertos	92.2%	98.31%
Menor tempo para identificação	5m33s	23s
Média aritmética de acertos	81%	98,68%
Média aritmética de tempo gasto	11m	1m11s

### 4.4. Discussão

Os resultados obtidos por meio da Rede Neural Convolutiva se provaram satisfatórios, obtendo valores acima de 90%. Quando comparado os resultados do Sistema Visual Humano com a Rede Neural Convolutiva desenvolvida tem-se que o tempo gasto para realizar a identificação por olho humano é, aproximadamente, 5 vezes maior que o necessário para o computador. Percebe-se uma economia de tempo, fator importante quando trata-se de uma grande quantidade de dados como na classificação de minerais. Além disso, deve-se levar em conta que este tempo gasto para a criação e desenvolvimento da RNA engloba a etapa de treino, sendo que o teste demanda um tempo menor.

Observa-se também que o contato com a atividade de classificação é um fator de grande impacto para o sucesso da identificação. Este fato torna-se evidente ao analisar e comparar os resultados obtidos por alunos de graduação em relação aos de pós-graduação. Além do voluntário 8 que, apesar de ser atualmente docente na área de geologia, obteve um número de acertos inferior ao de outros participantes. Além disso, ao observar as respostas fornecidas pelos usuários, percebeu-se que um dos grãos de Apatita não foi identificado corretamente por nenhum dos voluntários, sendo reconhecido pela RNA durante a fase de teste.

### 5. Considerações Finais

O processo de identificação de imagens de minerais possui alto custo para realização, em especial quando utilizamos o tempo como parâmetro. Assim, soluções inteligentes são necessárias como forma de garantir a produtividade e a qualidade do trabalho realizado por diversos profissionais. Neste estudo foi possível propor e treinar uma RNA Convolutiva capaz de identificar minerais através de imagens de microscópio. Os 3 modelos gerados alcançaram resultados satisfatórios com mais de 90% de acurácia no processo de reconhecimento e classificação de imagens de grãos de Apatita, Quartzo e Zircão. Os testes realizados com humanos revelaram que a RNA proposta possui capacidade para superar os esforços humanos na etapa de identificação, obtendo uma diferença de mais de 8 minutos no tempo gasto quando comparado com os voluntários. Isso mostra a capacidade da IA de obter e utilizar o conhecimento obtido por meio do treinamento para

diversos propósitos, como reconhecer e identificar minerais por meio de imagens. Para trabalhos futuros sugere-se a utilização de novas resoluções para criação dos Modelos de RNA, buscando aumentar a quantidade de características presentes nas imagens. Além disso, o aumento da base de dados original, tanto para treino quanto para teste, pode ser realizado buscando aumentar a quantidade de amostras disponíveis. Por fim, mudanças na arquitetura da RNA podem gerar novos resultados nos modelos, sendo possível melhorar sua performance desta forma.

## Referências

- Bayka, N. and Yilmaz, N. (2010). Mineral identification using color spaces and artificial neural networks. *Computers Geosciences*, 36.
- Borji, A. and Itti, L. (2014). Human vs. computer in scene and object recognition. *2014 IEEE Conference on Computer Vision and Pattern Recognition*.
- Bruand, E., Storey, C., and Fowler, M. (2014). Accessory mineral chemistry of high ba sr granites from northern scotland: Constraints on petrogenesis and records of whole-rock signature. *Journal of Petrology*, 55.
- Franco, A. O. B. (2006). Termocronologia por traços de fissão em apatitas na região do arco de ponta grossa, entre os alinhamentos de guapiara e são jeronimo-curiúva. Master's thesis, Universidade Estadual Paulista, Rio Claro Brasil.
- Izadi, H., Sadri, J., and Mahdokht, B. (2017). An intelligent system for mineral identification in thin sections based on a cascade approach. *Computers & Geoscience*.
- Jain, A. K., Mao, J., and Mohiuddin, K. (1996). Artificial neural networks: A tutorial. *Computer*, (3).
- Klein, C. and Dutrow, B. (2009). *Manual de ciência dos minerais*. Bookman Editora.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*.
- Maitre, J., Bouchard, K., and Bedard, P. (2019). Mineral grains recognition using computer vision and machine learning. *Computers Geosciences*, 27.
- Miynarczuk, M., Górszczyk, A., and Ślipek, B. (2013). The application of pattern recognition in the automatic clasffication on microscopic rock images. *Computers & Geoscience*.
- Pereira, R. and Porto, F. (2019). Deep learning application for plant classification on unbalanced training set. In *Anais do XIII Brazilian e-Science Workshop*. SBC.
- Thompson, S., Fueten, F., and Bockus, D. (2001). Mineral identification using artificial neural networks and the rotating polarizer stage. *Computers Geosciences*, 27.
- Tokui, S., Oono, K., Hido, S., and Clayton, J. (2015). Chainer: a next-generation open source framework for deep learning. In *Proceedings of workshop on machine learning systems (LearningSys) in the twenty-ninth annual conference on neural information processing systems (NIPS)*, pages 1–6.
- Zhang, M. (2017). *Raman Study of the Crystalline-to-Amorphous State in Alpha- Decay-Damaged Materials*.