

Classificação Automatizada de Rochas Ornamentais por Redes Neurais Convolucionais

Omitido para Revisão Cega

¹Omitido para Revisão Cega

Abstract. *Este trabalho propõe um sistema de Visão Computacional para classificação automatizada de rochas ornamentais a partir de imagens digitais. Utilizando transfer learning com a arquitetura ResNet18 e validação cruzada estratificada com 5 folds, o modelo foi treinado para identificar cinco classes de materiais pétreos. A metodologia empregou aumento de dados e busca em grade para otimização de hiperparâmetros, alcançando acurácia média de 98,62% ($\pm 0,83\%$) e coeficiente Kappa de 0,9826. Os resultados indicam a viabilidade da aplicação de Redes Neurais Convolucionais para discriminação de padrões visuais em geociências, contribuindo para a transformação digital do setor de rochas ornamentais.*

1. Introdução

O avanço das tecnologias de Inteligência Artificial (IA), em particular o *Deep Learning*, tem permitido sua aplicação em setores caracterizados por processos manuais e intensivos em conhecimento especializado, como a indústria de rochas ornamentais [Cheng et al. 2022]. O reconhecimento de padrões visuais, que antes dependia da experiência humana, pode agora ser automatizado por algoritmos de Visão Computacional. Este cenário tem impulsionado a adoção de ferramentas baseadas em IA para automatizar tarefas de inspeção e classificação visual em diversos domínios industriais.

A identificação de rochas ornamentais envolve a análise de texturas, cores, veios e inclusões, elementos que determinam sua aplicação estética e funcional, bem como seu valor comercial. Para profissionais não especializados, essa identificação a partir de fotografias é uma tarefa complexa e propensa a erros. A ausência de ferramentas acessíveis pode levar a escolhas inadequadas de materiais, impactando projetos de construção e design. Erros na especificação de materiais são uma das causas frequentes de retrabalho em obras residenciais e comerciais.

A ruptura do paradigma de análise visual manual ocorreu com a consolidação das Redes Neurais Convolucionais (CNNs), que permitem o aprendizado de características diretamente a partir dos dados [LeCun et al. 2015]. A capacidade de realizar extração automática de atributos elimina a necessidade de engenharia manual de características, processo que tradicionalmente exigia conhecimento especializado do domínio [Soares et al. 2025]. A evolução das arquiteturas de *Deep Learning* viabilizou também a execução desses modelos em dispositivos móveis, requisito relevante para inspeções em campo, onde o acesso a estações de trabalho dedicadas pode ser inviável.

O objetivo deste trabalho é desenvolver e validar um modelo de classificação de rochas ornamentais baseado em *transfer learning*, avaliando sua capacidade de discriminar cinco classes de materiais com características visuais similares. Adicionalmente, propõe-se a integração deste modelo em uma aplicação móvel que permita a identificação de rochas a partir de fotografias capturadas em campo.

2. Fundamentação Teórica

A digitalização tem se tornado estratégica no setor de construção e mineração [Kumar et al. 2023]. O segmento de rochas ornamentais possui importância econômica significativa nas indústrias de construção civil e design de interiores. O Brasil ocupa posição de destaque no cenário mundial, figurando entre os principais produtores e exportadores de granitos, mármore e quartzitos.

Apesar da maturidade industrial, o setor enfrenta desafios práticos de identificação. As rochas ornamentais apresentam variabilidade visual dentro de uma mesma classificação comercial e diferenças sutis entre materiais distintos, dependendo da mineralogia, tamanho e distribuição dos grãos, bem como da presença de estruturas secundárias. As dificuldades incluem a dependência de condições de iluminação, a avaliação subjetiva de textura e a interpretação de características microgeológicas visíveis a olho nu. A nomenclatura comercial frequentemente diverge da classificação geológica, gerando confusões que podem resultar em especificações incorretas de materiais.

O uso de IA para reconhecimento de padrões em materiais naturais evoluiu de soluções baseadas em processamento de imagens tradicionais, como histogramas e análise de cores, para abordagens automatizadas baseadas em aprendizado profundo [Liu et al. 2019]. Métodos tradicionais como SIFT (*Scale-Invariant Feature Transform*) e HOG (*Histogram of Oriented Gradients*), embora eficazes para objetos com geometria rígida, apresentam limitações ao lidar com a complexidade das superfícies naturais. A utilização de *transfer learning* consolidou-se como técnica eficaz para problemas com conjuntos de dados limitados [Garg and Dhiman 2021], permitindo aproveitar conhecimento aprendido em grandes bases de dados genéricas como o ImageNet.

A arquitetura ResNet (*Residual Network*) introduziu as conexões residuais (*skip connections*), que permitem o treinamento de redes profundas ao mitigar o problema do desvanecimento do gradiente [He et al. 2016]. A ResNet18, com suas 18 camadas, oferece equilíbrio entre capacidade de representação e custo computacional [Silva 2021]. Essa característica torna a arquitetura adequada para aplicações em dispositivos com recursos limitados, como smartphones e tablets utilizados em inspeções de campo.

A técnica de aumento de dados (*data augmentation*) consiste em aplicar transformações nas imagens de treino, como rotações, espelhamentos, recortes e ajustes de luminosidade, gerando variações que forçam a rede a aprender características invariantes às condições de captura [Shorten and Khoshgoftaar 2019]. Esta abordagem reduz o risco de sobreajuste (*overfitting*) e melhora a capacidade de generalização do modelo em cenários reais [Srivastava et al. 2014]. A combinação de *transfer learning* com aumento de dados tem se mostrado particularmente eficaz em domínios com escassez de dados rotulados.

3. Metodologia

A metodologia adotada segue um fluxo de desenvolvimento iterativo e multidisciplinar, conforme ilustrado na Figura 1, abrangendo desde a aquisição e o pré-processamento dos dados geológicos até o treinamento e a avaliação do modelo de *Deep Learning*.

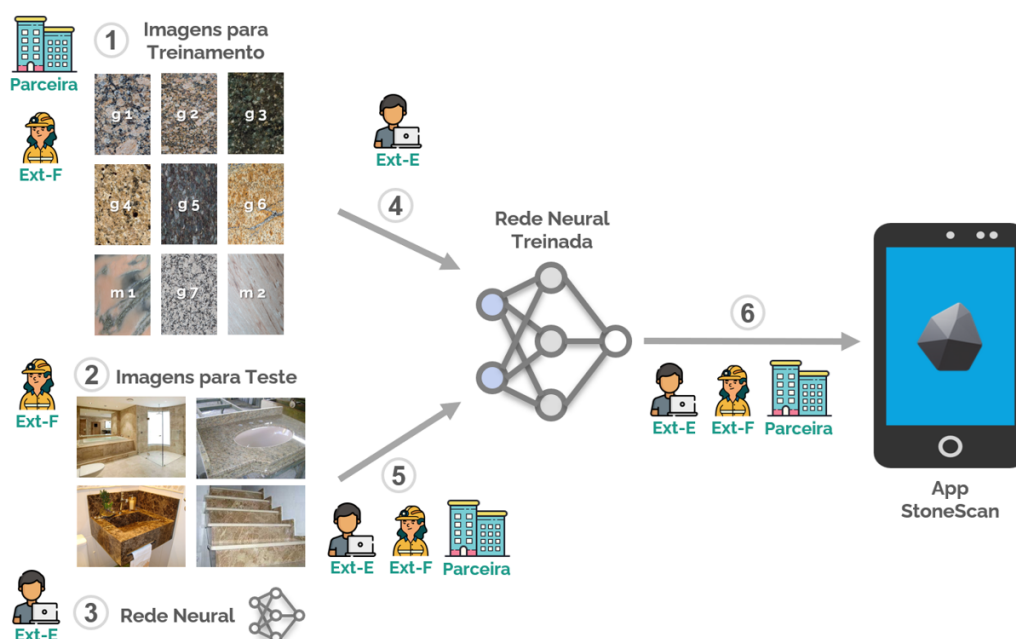


Figura 1. Fluxo metodológico do sistema proposto.

3.1. Conjunto de Dados

A etapa de coleta de imagens foi planejada para garantir diversidade e representatividade adequada do conjunto de dados. Foram capturadas aproximadamente 150 imagens para cada uma das cinco classes de materiais: Granito Branco Itaúnas, Mármore Matarazzo, Quartzito Perla, Quartzito Wakanda e Quartzito Verde Gaya. Cada material foi fotografado em diferentes distâncias, ângulos e condições de iluminação, utilizando dispositivos móveis.

Cabe ressaltar que o contexto industrial deste projeto contempla aproximadamente 20 classes diferentes de rochas ornamentais, as quais serão incorporadas progressivamente ao modelo até sua versão final. A seleção inicial de cinco classes deve-se ao fluxo de produção vigente, uma vez que a captura das imagens está condicionada à disponibilidade dos materiais durante o processo produtivo.

O protocolo de captura incluiu: imagens em regiões próximas para detalhes de textura; imagens em distância média para visão geral do padrão; imagens em ângulos inclinados para registrar efeitos de brilho e profundidade; e imagens com iluminação diferenciada (natural e artificial) para simular condições reais de observação. Este protocolo visa reduzir vies de captura e ampliar a robustez do modelo frente a variações encontradas em ambientes reais de uso.

A Figura 2 apresenta exemplos das classes utilizadas, evidenciando a variabilidade de padrões visuais e texturas características de cada material. Observa-se que, apesar de pertencerem a categorias geológicas distintas, alguns materiais apresentam similaridades visuais que tornam a tarefa de classificação desafiadora para observadores não especializados.



Figura 2. Exemplos das classes de rochas: (a) Granito Branco Itaúnas, (b) Mármore Matarazzo, (c) Quartzito Perla, (d) Quartzito Wakanda, (e) Quartzito Verde Gaya.

3.2. Arquitetura e Treinamento

Foi adotada a arquitetura ResNet18 com pesos pré-treinados no ImageNet, uma base de dados com mais de 14 milhões de imagens em 1.000 categorias. As camadas convolucionais foram congeladas para funcionar como extrator de características, aproveitando o conhecimento prévio da rede. A camada final foi substituída por camadas *dropout* (com taxas de 50% e 30%) e camadas totalmente conectadas para classificação das cinco classes. Esta estratégia reduz o número de parâmetros treináveis para aproximadamente 68.000, diminuindo o risco de sobreajuste e acelerando o treinamento.

Para avaliação do modelo, foi empregada validação cruzada estratificada com $k = 5$ partições. O pré-processamento incluiu técnicas de aumento de dados aplicadas apenas durante o treinamento: recortes aleatórios redimensionados, inversões horizontais e verticais, rotações de até 30° , ajustes de brilho, contraste e saturação, e desfoque gaussiano. É importante ressaltar que a avaliação foi realizada sobre as imagens originais, sem transformações de aumento, garantindo que as métricas reflitam o desempenho real do modelo.

A seleção de hiperparâmetros foi realizada por busca em grade (*Grid Search*), avaliando combinações de taxa de aprendizado (0.0001, 0.0005, 0.001), tamanho de lote (8, 16, 32) e otimizador (Adam, SGD, AdamW), totalizando 27 combinações. Cada combinação foi submetida à validação cruzada completa, resultando em 135 execuções de treinamento independentes. A Tabela 1 apresenta a configuração selecionada.

Tabela 1. Parâmetros do treinamento.

Parâmetro	Valor
Arquitetura	ResNet18 (pré-treinado)
Taxa de aprendizado	0.001
Tamanho do lote	16
Otimizador	Adam
Função de perda	Cross-Entropy (pesos de classe)
Validação cruzada	$k = 5$ folds
Early stopping	7 épocas
Tamanho da imagem	224×224 pixels

4. Resultados

A validação cruzada estratificada obteve acurácia média de 98,62% ($\pm 0,83\%$). A baixa variância entre os *folds* indica que o modelo apresenta desempenho consistente independentemente da partição utilizada para validação. A Tabela 2 apresenta os resultados por *fold*.

Tabela 2. Acurácia por *fold*.

Fold	Acurácia
1	98,11%
2	98,74%
3	98,74%
4	100,00%
5	97,48%
Média \pm DP	98,62% \pm 0,83%

A Tabela 3 apresenta precisão, revocação e F1-*score* por classe. Destaca-se que o Granito Branco Itaúnas obteve 100% de revocação (todas as amostras corretamente identificadas), enquanto o Quartzito Verde Gaya apresentou as métricas mais equilibradas.

Tabela 3. Métricas por classe.

Classe	Precisão	Revocação	F1	Suporte
Granito Branco Itaúnas	0,98	1,00	0,99	160
Mármore Matarazzo	0,99	0,97	0,98	156
Quartzito Perla	0,97	0,99	0,98	138
Quartzito Wakanda	0,99	0,98	0,99	134
Quartzito Verde Gaya	1,00	1,00	1,00	207
Média Ponderada	0,99	0,99	0,99	795

O coeficiente Kappa de Cohen foi de 0,9826, indicando concordância quase perfeita entre as predições do modelo e as classes verdadeiras. Este valor supera o limiar de 0,81 tipicamente considerado como concordância forte na literatura.

4.1. Análise de Convergência

A Figura 3 apresenta as curvas de aprendizado para os 5 *folds* da validação cruzada. A análise revela convergência consistente: as curvas de perda decrescem de forma suave, estabilizando-se nas épocas finais. O paralelismo entre as curvas de treino e validação indica ausência de sobreajuste significativo. Os valores de acurácia de validação atingem patamares superiores a 95% já nas primeiras épocas.

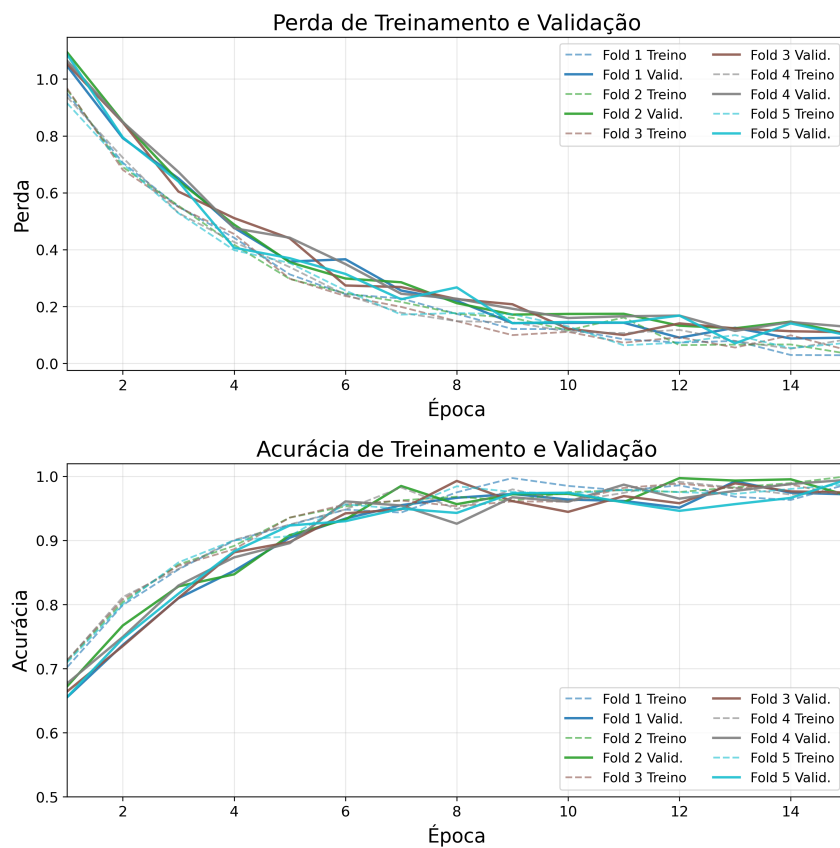


Figura 3. Curvas de aprendizado: (superior) perda; (inferior) acurácia. Linhas traçadas representam treinamento e linhas sólidas representam validação.

A Figura 4 apresenta a matriz de confusão agregada, evidenciando que a maioria das amostras foi classificada corretamente.

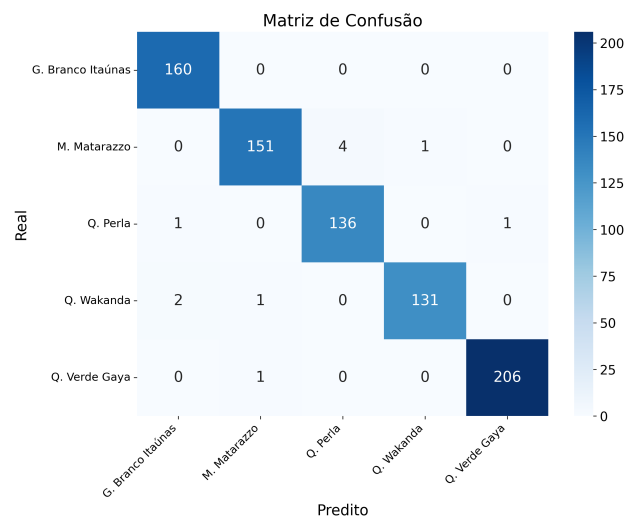


Figura 4. Matriz de confusão agregada dos 5 folds.

4.2. Discussão

A acurácia obtida pode ser justificada por diversos fatores metodológicos e características do domínio: (1) a validação cruzada estratificada garante resultados não dependentes de uma divisão específica dos dados; (2) múltiplas técnicas de regularização foram empregadas, incluindo *dropout*, *early stopping* e aumento de dados; (3) o congelamento das camadas convolucionais reduziu o número de parâmetros treináveis para aproximadamente 68.000; (4) rochas ornamentais apresentam padrões visuais distintos (granitos com grãos cristalinos, mármore com veios característicos, quartzitos com padrões cromáticos únicos) que facilitam a discriminação por CNNs.

Contudo, limitações devem ser reconhecidas: o desempenho está associado à qualidade do conjunto de dados, e o modelo pode requerer retreinamento para novos materiais.

5. Aplicação Prática

Do ponto de vista prático, o sistema visa democratizar o acesso à informação técnica, apoiando consumidores e profissionais na identificação de materiais. O aplicativo móvel desenvolvido integra o modelo de classificação a uma interface amigável, permitindo a identificação de rochas ornamentais diretamente em campo. A Figura 5 apresenta a interface do aplicativo, ilustrando o fluxo de uso.

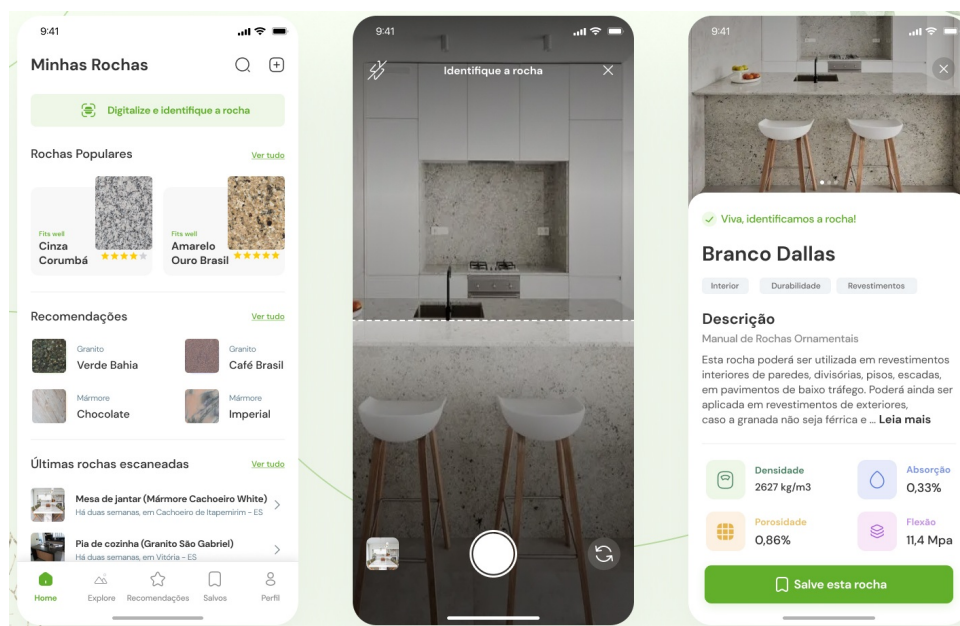


Figura 5. Interface do aplicativo: (a) tela inicial com histórico de escaneamentos, (b) interface de captura de imagem, (c) resultado com identificação da rocha e nível de confiança.

O fluxo de uso consiste em: (1) acesso à tela inicial com histórico de classificações anteriores; (2) captura de imagem da rocha via câmera do dispositivo; (3) processamento pelo modelo de classificação; (4) exibição do resultado com identificação da rocha e nível de confiança da predição. A arquitetura leve da ResNet18 viabiliza a execução em dispositivos móveis sem necessidade de conexão com servidores externos.

Como trabalho futuro, planeja-se a expansão do modelo para contemplar as aproximadamente 20 classes de rochas ornamentais do contexto industrial, além da coleta contínua de novas amostras para enriquecimento do conjunto de dados e melhoria da capacidade de generalização.

6. Considerações Finais

A contribuição deste trabalho reside na validação de que arquiteturas de *Deep Learning*, ajustadas via *transfer learning*, podem discriminar padrões visuais complexos em materiais geológicos. Os resultados obtidos (acurácia de 98,62% e Kappa de 0,9826) indicam a viabilidade técnica da abordagem para aplicações no setor de rochas ornamentais.

Trabalhos futuros incluem expansão do número de classes, coleta de dados em condições variadas, desenvolvimento do aplicativo móvel e implementação de ciclo MLOps.

Referências

- Cheng, L., Zhang, H., and Wang, Q. (2022). Deep learning applications in material science: a review. *Journal of Materials Research and Technology*, 21:1332–1350.
- Garg, S. and Dhiman, G. (2021). A comprehensive review of transfer learning for image classification. *Artificial Intelligence Review*, 54(1):555–605.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778.
- Kumar, S., Jain, P., and Singh, V. (2023). Digital transformation in the construction industry: A review of ai applications. In *Proceedings of the 2023 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, pages 1201–1207. IEEE.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.
- Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., and Pietikäinen, M. (2019). Deep learning for generic object detection: A survey. *International Journal of Computer Vision*, 128(2):261–318.
- Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):1–48.
- Silva, C. F. d. (2021). Redes residuais: Sobre quantidade de camadas e variância intraclasse utilizando aprendizagem não supervisionada. Dissertação de mestrado, Programa de Pós-Graduação em Informática, Universidade Federal da Paraíba (UFPB/CI), João Pessoa, PB.
- Soares, P., Ensina, L., and Foleis, J. (2025). A hybrid approach combining cnn and ensemble algorithms for dermoscopic image classification. In *Anais do XXII Encontro Nacional de Inteligência Artificial e Computacional*, pages 1527–1538, Porto Alegre, RS, Brasil. SBC.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958.