

INTELIGÊNCIA ARTIFICIAL PARA ANÁLISE DE SOLO COM BASE EM COR E TEXTURA

Emily Gabriela Tsen¹

Mariana Galvão Bispo¹

Renata Bruna dos Santos Cosco Favan²

João Ricardo Favan²

RESUMO

O estudo do solo é muito importante para um bom andamento da produção. Se iniciarmos a análise de um solo somente pela sua cor já é possível extrair informações extremamente valiosas sobre alguns atributos relevantes sobre o solo. O objetivo do projeto tem como finalidade desenvolver um modelo de Inteligência Artificial que baseada em uma rede neural que trouxesse para os agricultores uma forma mais fácil e integral de classificar as cores do solo entre preto e vermelho. Essa solução traz aos usuários a possibilidade de selecionar fotos do solo entre as tonalidades e receberem análises relacionadas à cor do solo detectado. Dessa forma, a plataforma ajuda agricultores e produtores a tomarem decisões mais rápidas sobre como manejar o solo, porém para ter uma avaliação completa, uma análise de laboratório ainda pode ser necessária.

Palavras-chave: Machine learning. Classificar. Solo.

1 INTRODUÇÃO

Nas últimas décadas, tem-se observado um aumento no consumo de alimentos e conseqüentemente no aumento da produção agrícola. Grover et al. (2024) nos traz a informação de que a saúde do solo desempenha um papel crucial na sustentabilidade da agricultura de nosso país.

Os pedologistas, fazem o estudo de cada solo através de técnicas visuais e laboratoriais, para dessa forma obter resultados sobre suas características e categorizá-lo. Mesmo que esses procedimentos sejam precisos, são lentos, custosos e dependem de avaliações subjetiva, acabando por resultar em falhas

¹ Discentes em Big Data no Agronegócio na FATEC Pompeia “Shunji Nishimura”, Pompéia - SP

² Docentes em Big Data no Agronegócio na FATEC Pompeia “Shunji Nishimura”, Pompeia - SP

humanas, especialmente na determinação precisa da cor do solo, que tem muitos desafios, tal qual a dificuldade em distinguir a tonalidade e gradações entre diferentes páginas de referência (Vilar, 2021). Além disso, o serviço é limitado quando se trata de áreas rurais distantes dos centros urbanos, pois faz o uso de equipamentos especiais e mão de obra capacitada. O estudo realizado por Park e Shin (2021), traz o uso de métodos computacionais, como o aprendizado de máquina, como alternativa para automatizar processos de análise de dados em diversos contextos. Essa estratégia também pode ser aplicada à classificação de solos, tornando-a mais acessível e eficiente.

Solos com coloração mais avermelhada são caracterizados por serem ricos em óxidos de ferro e outros nutrientes, como o manganês, cálcio e magnésio, por isso são recomendados para o plantio de culturas como cana-de-açúcar, soja e café. Entretanto, esse tipo de solo tem como principal propriedade a oxidação de ferro, que confirma a coloração avermelhada do solo (EMBRAPA, 2009). Estudos mostram que o cultivo correto pode alterar a forma e a distribuição dos óxidos de ferro, impactando a fertilidade do solo, como retratados em pesquisas realizadas no noroeste do Irã (REZAPOUR et al., 2015). Quando corrigido, esse tipo de solo se torna altamente apropriado para a produção de culturas de extrema importância no Brasil.

O solo com coloração preta, com altos teores de matéria orgânica e capacidade de reter água, é apropriado para a produção de hortaliças e grãos, como soja e arroz. A capacidade de acumular água e nutrientes torna este solo perfeito para a agricultura e outras culturas que exigem maior estabilidade hídrica. Estudos indicam que os solos orgânicos possuem altos níveis de nutrientes essenciais, como nitrogênio, fósforo e potássio, além de ter uma estrutura solta que evita a compactação e facilita a retenção de água (BOUCHER, 2023). Uma boa avaliação dos tipos de solo pode proporcionar percepções valiosas tanto para grandes agricultores quanto para indivíduos com pequenas propriedades rurais.

Os valores de uma análise de solo no Brasil podem chegar a R\$500, quando se trata de uma vasta área para análise, podendo variar para mais ou para menos dependendo da sua complexidade (Universidade Federal do Piauí, 2018). Os custos podem aumentar rapidamente em grandes áreas agrícolas, onde podem ser

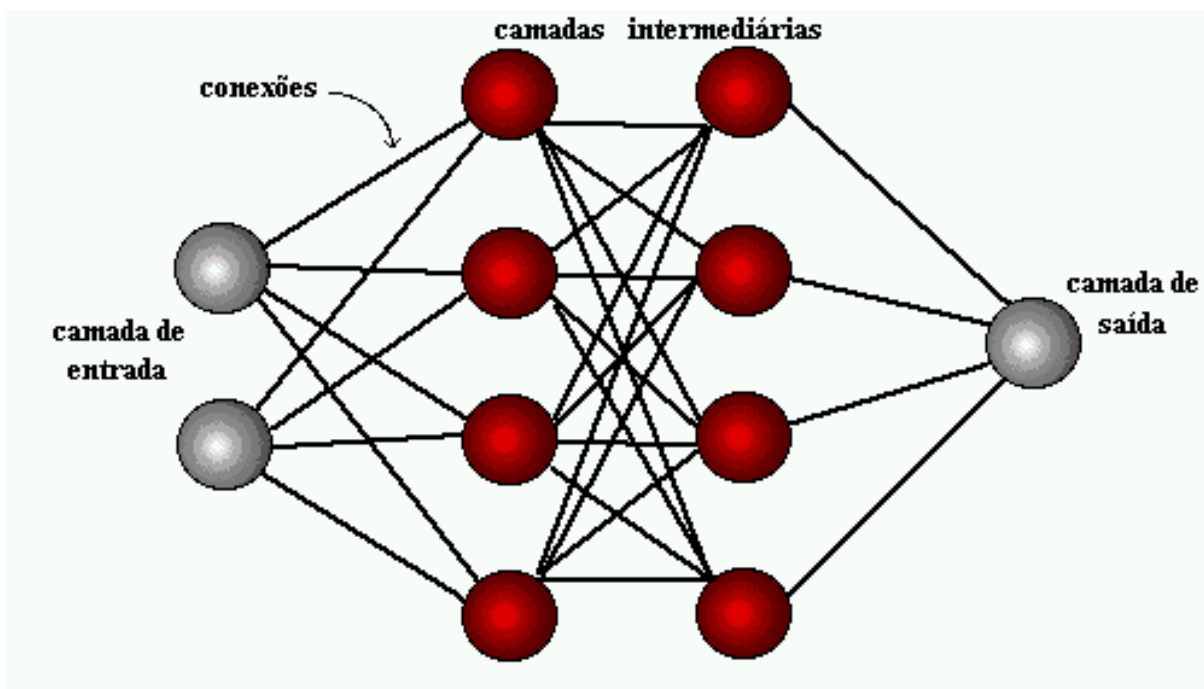
necessárias centenas ou milhares de amostras. Por causa do alto custo, em alguns casos os produtores limitam os seus espaços para adequar a um número de amostras o que pode acabar não representando de forma integral a fertilidade da área total. A coleta e envio de amostras para laboratórios especializados, especialmente em locais remotos, pode levar a custos adicionais.

Ao fazer uma pesquisa de campo no Brasil, foi possível perceber que os laboratórios de solos localizados no estado de São Paulo, têm grande relevância em pesquisas nacionais e internacionais, segundo o Instituto Agrônomo de Campinas (IAC) (2024). Isso faz com que muitos pequenos agricultores e pessoas que vivem em cidades rurais em outros estados não possam seguir o processo tradicional devido aos custos, falta de acesso a equipamentos, profissionais qualificados e o tempo necessário para análise. Ao fornecer um método mais rápido e imediato para a classificação e análise do solo, uma implementação baseada em inteligência artificial pode reduzir significativamente esses problemas.

As Redes Neurais Convolucionais (CNN) são uma classe de modelos de aprendizado de máquina amplamente utilizadas na análise de imagens devido à sua capacidade de identificar padrões complexos. Inspiradas no funcionamento do cérebro humano, essas redes são compostas por camadas de neurônios artificiais, onde cada camada extrai características específicas dos dados de entrada, como texturas, formas e cores (Ferneda, 2003).

A rede utilizada no projeto é uma peça fundamental para a classificação das imagens de solo. O processo inicia na camada de entrada, onde as imagens do solo são inseridas. Em seguida, as camadas intermediárias ajustam os pesos com base no erro estimado em cada previsão, processo conhecido como *backpropagation*. A camada de saída fornece a previsão final, indicando a probabilidade de o solo ser vermelho ou preto. A Figura 1 ilustra visualmente essa estrutura de rede neural, mostrando como as camadas processam as informações para fornecer a previsão final, que indica a classe do solo.

Figura 1: Ilustração de uma rede neural.



Fonte: USP. Neural Networks Research.

Esses dados gerados são essenciais para os agricultores e especialistas do ramo, pois possibilitam uma melhor compreensão das propriedades do solo, respondendo diretamente questões que podem ser duvidosas durante a atividade agrícola. A utilização dessa inovação representa um progresso, considerando uma avaliação mais ágil, exata e acessível na agricultura contemporânea.

O propósito do projeto é utilizar a tecnologia de aprendizado de máquina para reconhecer e classificar o solo em duas categorias principais: vermelho e preto, exclusivamente com base em imagens do solo, fornecendo uma probabilidade percentual para cada uma dessas categorias.

2 MATERIAL E MÉTODOS

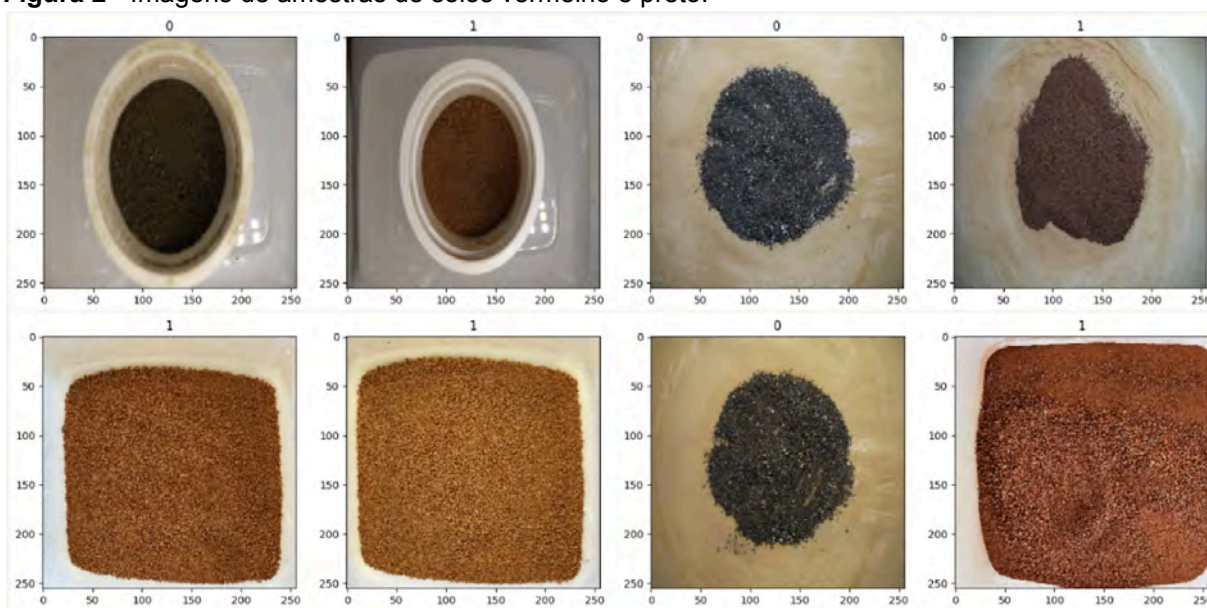
Coleta de Dados

Para este estudo, foram coletadas no total 400 imagens de solo, sendo que 200 imagens foram de solo vermelho e as outras 200, foram de solo preto. O local de coleta foi um Laboratório de Solos da Fundação Shunji Nishimura, localizado em Pompeia, SP. Os solos coletados para as fotografias foram pegos diretamente dos

mostruários do laboratório, no qual representavam amostras de diferentes regiões geográficas, para garantir diferentes tons e granulometrias de cada tipo de solo. As fotografias foram realizadas em alta definição, com uma câmera profissional de 16 megapixels, dentro de uma caixa branca iluminada, o que garantiu a captação precisa das características dos solos. No decorrer da captura, foram feitos ajustes no zoom e na iluminação, mudando a luz de branca para amarela. Para melhorar a qualidade das fotografias, desenvolveu-se uma caixa de papelão revestida com papel branco, simulando um pequeno estúdio fotográfico. Contudo, este pequeno estúdio improvisado não vedava totalmente a luz natural, o que pode ter provocado ligeiras variações nas condições de iluminação. Adicionalmente, a forma dos recipientes foi alterada e os grãos foram continuamente agitados para assegurar a captura de diversas variações dos solos. Utilizou-se uma câmera profissional para garantir maior exatidão e qualidade nas fotografias obtidas. Foi preservado um ambiente controlado para prevenir interferências externas, garantindo a consistência dos dados coletados.

A Figura 2 traz uma ilustração dos exemplos das amostras de solo coletadas, apresentando as diferenças visuais entre o solo vermelho e o solo preto.

Figura 2 - Imagens de amostras de solos vermelho e preto.



Fonte: Elaborada pelo autor (2024).

Pré-processamento

Tendo em vista o modelo de aprendizado de máquina que seria usado, as fotos coletadas foram inseridas em um pré-processamento para atender os requisitos principais. Foram reduzidas todas as imagens para um tamanho fixo e convertidas para escala de cinza, com o objetivo de facilitar o processamento e a análise com redes neurais, similar ao método de reconhecimento de variedades descrito por Khatchourian e Padilha (2008), onde filtros de *anti-aliasing* e técnicas de segmentação foram aplicados para destacar características visuais específicas. Além disso, os valores de pixel foram ajustados para intervalos de 0 a 1, o que trouxe uma grande melhoria para os resultados de performance da rede neural durante o treinamento (Ultralytics, 2024).

Após a fase dessas conversões, houve a divisão do conjunto de dados em três partes: 70% das imagens foram direcionadas para treinamento, 20% para validação e 10% para teste (Komesu, 2022). Também, essa metodologia foi baseada em estudos que mostraram que para trazer melhores resultados e precisão do modelo é necessário usar de 70 a 80% dos dados para treinamento e o uso de 20 a 30% para validação e teste (Gholamy; Kreinovich; Kosheleva, 2018).

Ajuste do Modelo

Para a formação do modelo em python definimos algumas fases iniciais. A rede neural convolucional foi estruturada em três tipos principais de camadas, as Camadas Convolucionais que são responsáveis por identificar características importantes nas imagens, como padrões e texturas. As Camadas de Pooling, que servem para "reduzir" a imagem, mantendo apenas os aspectos mais importantes. E as Camadas Densas que integram todas as informações identificadas anteriormente, facilitando a decisão final.

Para tornar as operações mais eficientes, foi usada uma função chamada ReLU nas camadas convolucionais e densas, ajudando o modelo a aprender de forma mais rápida e precisa. Na última camada, também se aplicou uma função *sigmoid*, que funciona como um "interruptor", permitindo que o modelo classifique a imagem em uma das duas categorias de solo (vermelho ou preto). Também foi empregada uma técnica de regularização *dropout* nas camadas densas para fazer uma performance melhor e evitar o *overfitting*.

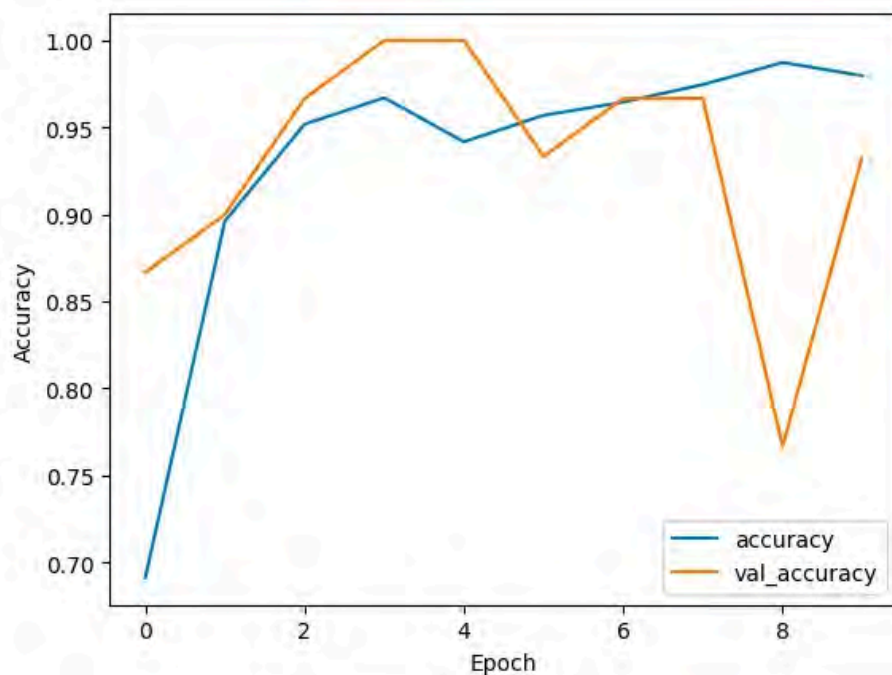
O modelo foi treinado usando o otimizador Adam e a função de perda binary crossentropy, bem adaptados para problemas de classificação binária. Todo o desenvolvimento foi realizado com as bibliotecas TensorFlow 2.16.1 (Google Brain, 2024) e Keras 3.0.5 (Chollet, 2024), que fornecem ferramentas avançadas para a construção e treinamento de redes neurais profundas (Abadi et al., 2015).

Validação do Modelo

Para avaliar o desempenho do modelo de rede neural convolucional, utilizou-se métricas de acurácia, precisão, revocação e F1-score, que são indicadores fundamentais para validação de modelos de classificação binária (Müller; Guido, 2016; Brownlee, 2019). A validação foi realizada com 20% dos dados reservados para essa finalidade e monitorada ao longo de 10 épocas. A acurácia representa a taxa de acertos gerais do modelo, enquanto a precisão, revocação e F1-score fornecem uma análise mais detalhada sobre a capacidade do modelo em identificar corretamente cada classe, minimizando falsos positivos e falsos negativos (Santos et al., 2016; Saito; Rehmsmeier, 2015).

A Curva de Acurácia (Figura 4) é um tipo de gráfico que mostra como o desempenho do modelo evolui à medida que ele é treinado, permitindo a comparação entre a acurácia no conjunto de dados de treinamento (linha azul) e no conjunto de validação (linha laranja). A linha azul indica a acurácia do modelo sobre os dados de treinamento em cada época, enquanto a linha laranja monitora a acurácia do modelo nos dados de validação, que o modelo não viu durante o treinamento.

Figura 4: Gráfico referente às curvas e validação da acurácia ao longo das épocas de treinamento de um modelo de aprendizado de máquina.



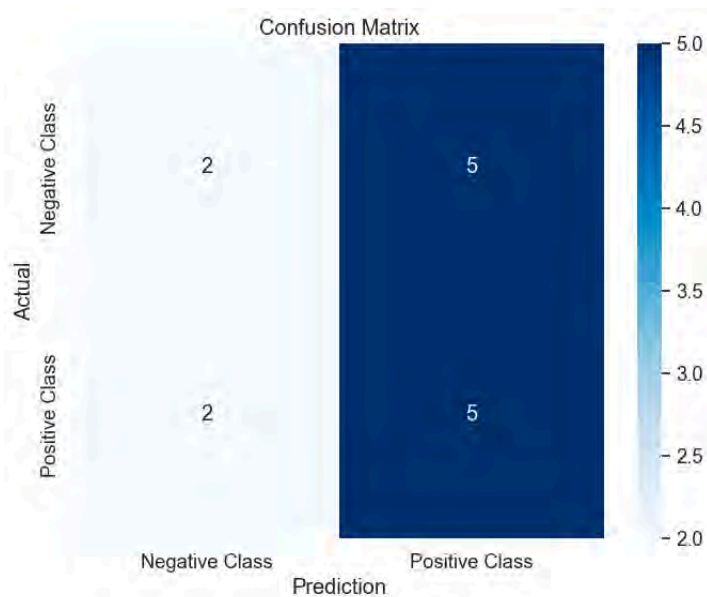
Fonte: Elaborada pelo autor (2024).

Este método de validação foi crucial para avaliar a capacidade do modelo de generalizar, assegurando que ele aprenda os padrões nos dados de treinamento e seja capaz de fazer previsões precisas em dados novos (GOODFELLOW *et al.*, 2016). Durante o treinamento, é fundamental monitorar continuamente o desempenho para identificar problemas como o *overfitting*, que ocorre quando o modelo ajusta-se excessivamente aos dados de treinamento, mas apresenta dificuldades em generalizar para novos dados (BISHOP, 2006). Ajustando os hiperparâmetros com base no desempenho no conjunto de validação, foi possível

refinar o modelo e alcançar um equilíbrio entre aprendizado eficaz e generalização (GERON, 2019).

Notou-se que foi preciso refazer o treinamento mais uma vez, pois em alguns casos o modelo passava a decorar os dados e não estava trazendo um resultado esperado de inteligência. Pois, por mais que a amostra dos dados possam ser relativamente “pequenas” para outros projetos, para a rede usada nesta pesquisa era significativamente suficiente, não precisando mais que 10 épocas de treinamento. Sendo assim, até chegar nessa conclusão, foi necessário repetir algumas vezes o treinamento para chegar ao equilíbrio exato fazendo com que a IA chegasse à definição exata dos solos sem ter memorizado eles.

Figura 5: A matriz de confusão gerada como forma visual de avaliar o desempenho de um modelo de classificação.



Fonte: Elaborada pelo autor (2024).

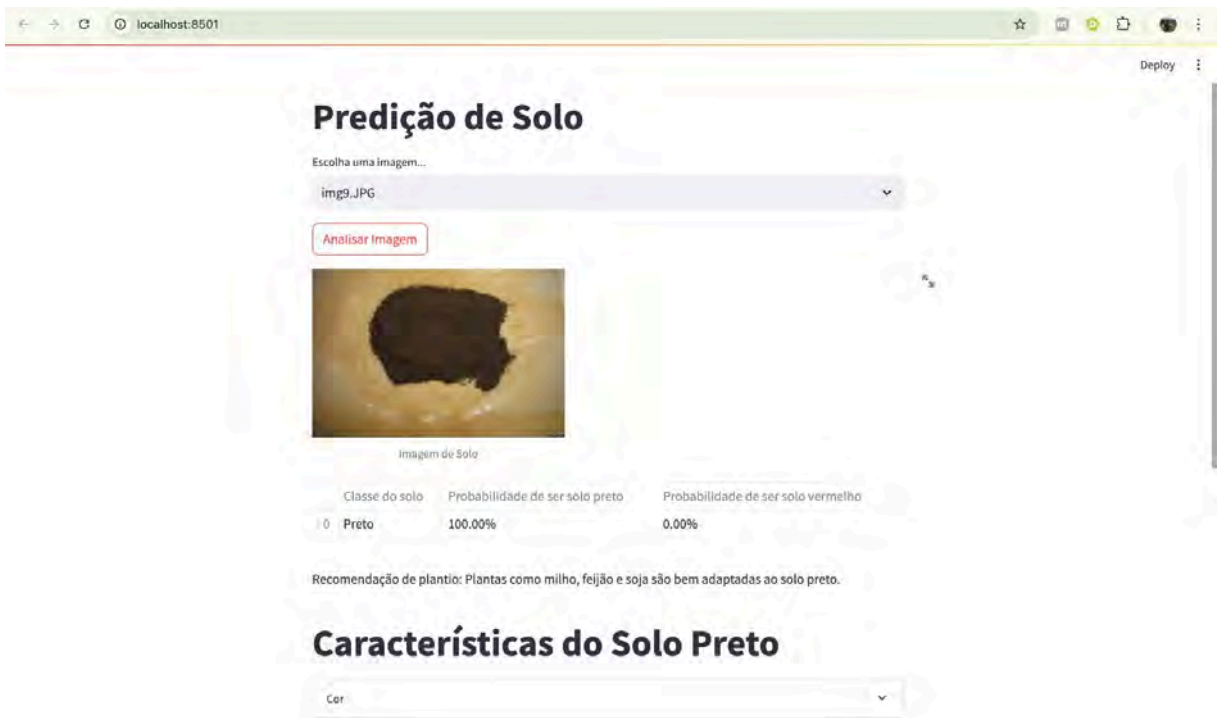
No eixo vertical da matriz de confusão (Figura 5), estão representadas as classes reais (também chamadas de valores "esperados"), enquanto no eixo horizontal estão as classes previstas pelo modelo. A análise da matriz mostrou que o modelo conseguiu prever corretamente 2 amostras da classe negativa (solo preto) e 5 amostras da classe positiva (solo vermelho). Entretanto, o modelo cometeu erros em 5 casos, onde classificou incorretamente amostras da classe negativa como positiva, e em 2 casos em que amostras da classe positiva foram classificadas como

negativas.

3 RESULTADOS E DISCUSSÃO

A ferramenta desenvolvida foi capaz de classificar as imagens de solo com uma acurácia de até 95%, diferenciando de forma eficiente entre solos. Utilizando uma CNN, o critério principal para a classificação foi a identificação precisa da cor do solo. No entanto, devido à limitação do modelo em classificar apenas as categorias "vermelho" e "preto", algumas tonalidades intermediárias, como variações de marrom, podem causar incertezas na classificação. Nessas situações, a IA apresenta probabilidades divididas, refletindo a dificuldade em determinar com precisão a cor predominante do solo.

Figura 6: Primeira parte da tela de aplicação do sistema de previsão.



Fonte: Elaborada pelo autor (2024).

Na Figura 6, é possível observar uma interface responsiva criada com o Streamlit (um *work frame* de desenvolvimento em python) para possibilitar a utilização do modelo de IA por usuários não técnicos (Treuille e Kelly, 2023). A interface apresenta um campo de input de imagens (no qual colocamos os 10% das

imagens de teste) entre tons variados de ambas as colorações para que o usuário possa escolher de sua preferência. Também, ao escolher a imagem o modelo é processado e apresenta-se a probabilidade da classe do solo juntamente com uma sugestão de plantio, indicando quais espécies são mais apropriadas para o tipo de solo detectado.

A inteligência artificial, utilizando imagens do solo, analisa e categoriza o tipo de solo em vermelho ou preto, com base nas características visuais extraídas das imagens. Apesar do modelo de Inteligência Artificial ser encarregado apenas da classificação das imagens, a plataforma amplia os resultados ao fornecer informações extras sobre o solo, tais como cor, textura, composição mineral, fertilidade, nível de oxidação, resistência à erosão e localização geográfica. Essas informações foram integradas ao sistema com base em pesquisas e estudos anteriores. Como demonstrada na Figura 7, a plataforma também recomenda a variedade de plantio mais apropriada para cada tipo de solo.

Figura 7: Segunda parte da tela de aplicação do sistema de previsão.



Fonte: Elaborada pelo autor (2024).

Esses dados extras não são produzidos pelo modelo de Inteligência Artificial, mas são exibidos com base na classificação feita, empregando uma lógica condicional. Portanto, a plataforma proporciona aos produtores rurais um recurso eficaz e econômico para aprimorar suas práticas agrícolas. A solução facilita a gestão do solo promovendo a sustentabilidade, auxiliando no crescimento da produtividade e na redução do efeito no meio ambiente.

A partir da classificação da cor do solo, o sistema foi capaz de sugerir práticas agrícolas adequadas para cada tipo, permitindo que tanto agricultores quanto cultivadores de hortas domésticas aprimorassem suas colheitas com base nas informações associadas.

Ao longo do treinamento do modelo, os indicadores de acurácia e perda apontaram um progresso notável na habilidade de classificar. Depois de 10 períodos, o modelo atingiu a estabilidade, exibindo perdas mínimas e um crescimento constante na precisão. Isso indica que as fases de pré-processamento e ajuste de hiperparâmetros foram eficientes para prevenir o superajuste e assegurar uma boa generalização.

A ferramenta também se mostrou promissora em termos de acessibilidade. Utilizando uma plataforma simples de upload de imagens, como o Streamlit, a solução se torna mais disponível para uma ampla gama de usuários, eliminando a necessidade de equipamentos sofisticados para a análise de solo. Essa abordagem democratiza o acesso a dados precisos, tornando a classificação de solos uma atividade mais acessível e economicamente viável.

4 CONSIDERAÇÕES FINAIS

Os achados sugerem que a abordagem sugerida tem um grande potencial para aprimorar as práticas agrícolas, oferecendo uma opção prática e econômica para a identificação inicial de atributos do solo, como sua fertilidade e habilidade de reter umidade. Contudo, notou-se a exigência de validar o modelo em uma ampla gama de tipos de solo e condições ambientais, um obstáculo que limita a utilização da ferramenta além dos cenários analisados neste estudo.

Para progressos futuros, sugere-se expandir a base de dados com amostras que incluem uma variedade mais ampla de tipos de solo e condições climáticas. Isso ajudará a tornar o modelo mais sólido e de fácil aplicação. Adicionalmente, a adição de tecnologias adicionais, tais como sensores de umidade e instrumentos de análise de imagens de alta definição, pode melhorar a exatidão da solução. Esta melhoria possibilitará que o software realize diagnósticos ainda mais minuciosos, promovendo práticas de agricultura de precisão. Com essas melhorias, o modelo poderá se estabelecer como um instrumento crucial para uma administração agrícola mais eficaz e sustentável, afetando de forma positiva tanto a produtividade quanto a preservação dos recursos naturais.

REFERÊNCIAS

AGROADVANCE. **Análise de solo: O que é, tipos e como fazer corretamente.** 2024. Disponível em: <https://agroadvance.com.br/blog-analise-de-solo-o-que-e-tipos-e-como-fazer/>.

Acesso em: 24 out. 2024.

BAYER. **Como realizar uma boa análise de solo.** 2024. Disponível em: <https://www.agro.bayer.com.br/conteudos-impulso-bayer/como-realizar-uma-bo-a-analise-de-solo>. Acesso em: 24 out. 2024.

BOUCHER, Ethel. **Is Black Earth Good For Vegetable Gardens?**. Vegetable Gardening News, 2023. Disponível em: https://www.vegetablegardeningnews.com/is-black-earth-good-for-vegetable-gardens/#google_vignette . Acesso em: 7 out. 2024.

DEMATTE, J.; DOTTO, A.; BEDIN, L.; SAYÃO, V.; SOUZA, A. **Soil analytical quality control by traditional and spectroscopy techniques: Constructing the future of a hybrid laboratory for low environmental impact.** Geoderma, 2019. Disponível em: <https://doi.org/10.1016/J.GEODERMA.2018.09.010>. Acesso em: 30 set. 2024.

EMBRAPA. **A cor do solo: interpretando as cores do solo.** Manaus: Embrapa Amazônia Ocidental, 2009. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/195720/1/A-Cor-Do-Solo-Interpretando-as-Cores-Do.pdf>. Acesso em: 21 out. 2024.

EMBRAPA. **Solos do Brasil.** Empresa Brasileira de Pesquisa Agropecuária, 2020. Disponível em: <https://www.embrapa.br/tema-solos-brasileiros/solos-do-brasil>. Acesso em: 22 out. 2024.

FERNEDA, E. **Redes neurais e sua aplicação em sistemas de recuperação de informação.** Ciência da Informação, v. 32, n. 3, p. 27-36, 2003. Disponível em: <https://www.scielo.br/j/ci/a/SQ9myjZWLxnyXfstXMgCdch>. Acesso em: 24 out. 2024.

GHOLAMY, A.; KREINOVICH, V.; KOSHELEVA, O. **Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation.** El Paso: University of Texas at El Paso, 2018. Disponível em: https://scholarworks.utep.edu/cs_techrep/1209/. Acesso em: 7 out. 2024.

GROVER, D.; MISHRA, A.K.; RANI, P.; KALONIA, N.; CHAUDHARY, A.; SHARMA, S. **Soil Management in Sustainable Agriculture: Principles and Techniques.** In: KUMAR, P.; AISHWARYA, (org.). **Technological Approaches for Climate Smart Agriculture.** Cham: Springer, 2024. Disponível em: https://doi.org/10.1007/978-3-031-52708-1_4. Acesso em: 30 set. 2024.

KHATCHATOURIAN, O.; PADILHA, F. R. R. **Reconhecimento de variedades de soja por meio do processamento de imagens digitais usando redes neurais artificiais.** Engenharia Agrícola, Jaboticabal, v. 28, n. 4, p. 759-769, 2008.

Disponível em: <https://www.scielo.br/j/eagri/a/F5kVxxcfHKDcDKnHKF7NKkv/>. Acesso em: 24 out. 2024.

KOMESU, D. K. **Divisão de dados em Treino, Validação e Teste para Machine Learning**. 15 fev. 2022. Disponível em: <https://dkko.me/posts/treino-teste-validacao/>. Acesso em: 24 out. 2024.

PARK, Y.; SHIN, Y.-W. **Trend Analysis of Pet Plants Before and After COVID-19 Outbreak Using Topic Modeling: Focusing on Big Data of News Articles from 2018 to 2021**. 30 de dez. de 2021. Disponível em: <https://jppe.ppe.or.kr/journal/view.php?doi=10.11628/ksppe.2021.24.6.563>. Acesso em: 7 out. 2024.

REZAPOUR, S.; AZHAH, H.; MOMTAZ, H.R. et al. **Changes in forms and distribution pattern of soil iron oxides due to long-term cropping in the Northwest of Iran**. *Environmental Earth Sciences*, v. 73, p. 7275–7286, 2015. Disponível em: <https://doi.org/10.1007/s12665-014-3933-y>. Acesso em: 09 out. 2024.

TENSORFLOW. *CNN image classification tutorial*. 2024. Disponível em: <https://www.tensorflow.org/tutorials/images/cnn?hl=pt-br>. Acesso em: 24 out. 2024.

TREUILLE, Adrien; KELLY, Amanda. **Generative AI and Streamlit: A perfect match**. Streamlit Blog, 15 jun. 2023. Disponível em: <https://blog.streamlit.io/generative-ai-and-streamlit-a-perfect-match/>. Acesso em: 24 out. 2024.

UNIVERSIDADE FEDERAL DE SANTA CATARINA. *Métricas de desempenho para classificação*. 2020. Disponível em: <https://geam.paginas.ufsc.br/files/2020/02/Metricas-desempenho-classificacao2-1.pdf>. Acesso em: 20 out. 2024.

UNIVERSIDADE FEDERAL DO PIAUÍ (UFPI). **Tabela 3 – Preços da safra 2017-2018**. Disponível em: https://www.ufpi.br/arquivos_download/arquivos/CPCE/Tabela_3-precos_safra_2017-2018.pdf. Acesso em: 24 out. 2024.

ULTRALYTICS. **Data Preprocessing Techniques for Annotated Computer Vision Data**. 2024. Disponível em: https://docs.ultralytics.com/pt/guides/preprocessing_annotated_data/. Acesso em: 20 out. 2024.

USP. **Neural Networks Research**. Instituto de Ciências Matemáticas e de Computação. Disponível em: <https://sites.icmc.usp.br/andre/research/neural/>. Acesso em: 7 out. 2024.

VILAR, Daniel. **Descobrimo as cores do solo na prática**. 9 abr. 2021. Disponível em: <https://www.agriconline.com.br/portal/artigo/descobrimo-as-cores-do-solo-na-pratica/>. Acesso em: 18 out. 2024.

INSTITUTO AGRONÔMICO DE CAMPINAS (IAC). Laboratórios de análise de solo e planta são fundamentais para a agricultura sustentável. Disponível em: <https://www.iac.sp.gov.br/noticiasdetalhes.php?tag=1158>. Acesso em: 16 out. 2024.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. Springer, 2006.

GERON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow**. 2. ed. Sebastopol: O'Reilly Media, 2019.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge: MIT Press, 2016.

MÜLLER, Andreas C.; GUIDO, Sarah. **Introduction to Machine Learning with Python: A Guide for Data Scientists**. 1. ed. Sebastopol: O'Reilly Media, 2016.

BROWNLEE, Jason. **Classification Accuracy, Precision, Recall, F-Score: What's the Difference?** 14 jul. 2019. Disponível em: <https://machinelearningmastery.com/classification-accuracy-precision-recall-f-score/>. Acesso em: 18 out. 2024.

SAITO, T.; REHMSMEIER, M. **The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets**. 2015. Disponível em: <https://www.jmlr.org/papers/volume18/15-535/15-535.pdf>. Acesso em: 18 out. 2024.

SANTOS, L. S. et al. **A Performance Analysis of Machine Learning Classifiers for Imbalanced Datasets**. 2016. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877050916301611>. Acesso em: 18 out. 2024.