

APLICAÇÃO DE ALGORITMOS DE APRENDIZADO DE MÁQUINA NA CLASSIFICAÇÃO TEXTURAL DE SOLOS ATRAVÉS DA ANÁLISE DE IMAGENS

Gabriel Carneiro Aleixo¹

gabriel.aleixo2@fatec.sp.gov.br

Fatec Pompeia Shunji Nishimura

João Ricardo Favan

Fatec Pompeia Shunji Nishimura

Gustavo Di Chiacchio Faulin

gustavo.faulin@fatec.sp.gov.br

Fatec Pompeia Shunji Nishimura

1. INTRODUÇÃO

A textura do solo é um atributo físico fundamental para a agricultura, pois influencia diretamente a retenção de água, a disponibilidade de nutrientes e, consequentemente, a produtividade das culturas. Tradicionalmente, sua determinação é realizada por métodos laboratoriais, como a análise granulométrica por peneiramento e sedimentação, que, embora precisos, demandam tempo, dependem de infraestrutura especializada e geram custos elevados. Nesse contexto, este estudo investigou a viabilidade de um método mais ágil e acessível, utilizando imagens digitais e algoritmos de aprendizado de máquina. A relevância desta pesquisa justifica-se pela busca por otimização na tomada de decisões agrícolas, alinhada aos princípios da agricultura de precisão. O objetivo principal foi desenvolver um algoritmo capaz de classificar a textura do solo com base em imagens obtidas por microscópio digital e *smartphones*, comparando a acurácia dos diferentes dispositivos.

2. METODOLOGIA

Para a formação do banco de dados, foram selecionadas amostras de solo provenientes de

diversas regiões do Brasil, submetidas à análise granulométrica pelo Método da Pipeta, conforme os procedimentos descritos no *Manual de Métodos de Análise de Solo* (DONAGEMMA et al., 2017) [1]. As análises foram realizadas no Laboratório de Solos e Tecido Vegetal da Fundação Shunji Nishimura de Tecnologia, instituição parceira desta pesquisa.

O foco do estudo foi direcionado às quatro classes de maior representatividade em solos agrícolas brasileiros: arenoso, médio, argiloso e muito argiloso, as quais apresentam diferenças relevantes na dinâmica de água e nutrientes no perfil do solo.

As imagens das amostras foram capturadas utilizando três protocolos padronizados de captura, projetados para avaliar o impacto da iluminação e do dispositivo de aquisição sobre a qualidade das imagens e o desempenho dos algoritmos:

- Microscópio Digital MOYSUWE MDM201 Pro (16 MP, 1300×), para imagens de alta resolução;
- Smartphone com Iluminação Direta, utilizando suporte e *Ring Light* controlada em ambiente escuro;
- Smartphone com Iluminação Difusa, utilizando uma caixa revestida internamente com papel alumínio para criar iluminação uniforme.

As imagens obtidas foram processadas e analisadas na plataforma Orange Data Mining

(versão 3.37). Foram testados seis algoritmos de aprendizado de máquina supervisionado (Regressão Logística, SVM, k-NN, Random Forest, Árvore de Decisão e Rede Neural) com o objetivo de identificar aquele com melhor desempenho na classificação textural do solo.

A avaliação dos modelos foi realizada por meio de validação cruzada com 10 *folds*, garantindo maior robustez estatística e controle sobre o sobreajuste (*overfitting*). As métricas de desempenho utilizadas foram Acurácia de Classificação (CA), Precisão e F1-Score.

3. RESULTADOS E DISCUSSÕES

Os resultados obtidos confirmam a alta viabilidade do uso de aprendizado de máquina na classificação textural de solos a partir de imagens digitais. A análise comparativa entre os protocolos de captura e os algoritmos avaliados demonstrou que é possível alcançar elevada acurácia e consistência utilizando dispositivos de baixo custo, como smartphones, desde que a padronização da iluminação e da captura seja rigorosamente controlada.

Os resultados médios de Acurácia de Classificação (CA) estão apresentados na Tabela 1.

Tabela 1 – Acurácia média dos algoritmos em diferentes métodos de captura.

Algoritmo	Microscópio Digital	Smartphone (luz direta)	Smartphone (luz difusa)
Regressão Logística	91,1%	97,5%	81,2%
SVM	84,4%	85,0%	77,5%
k-NN	81,1%	79,8%	74,1%
Random Forest	74,3%	73,4%	75,3%
Tree	68,8%	76,9%	76,2%
Neural Network	67,9%	65,9%	66,9%

Fonte: Autoria própria (2025).

Desempenho geral dos algoritmos

Em todos os cenários, a Regressão Logística foi o modelo mais estável e preciso, superando os demais algoritmos em até 25

pontos percentuais. Isso demonstra que, para conjuntos de dados com classes bem definidas e número moderado de amostras, modelos lineares apresentam maior capacidade de generalização e menor sensibilidade a ruídos de iluminação ou textura.

O algoritmo SVM obteve resultados consistentes, porém inferiores, o que sugere que a separação entre as classes texturais pode ser representada de forma satisfatória por fronteiras lineares. Já os modelos baseados em árvores (*Random Forest e Tree*) apresentaram desempenho inferior e maior variabilidade, evidenciando sensibilidade a pequenas diferenças no padrão das partículas de solo.

Influência do método de captura

A comparação entre os três métodos de captura revelou uma hierarquia de desempenho bem definida. O protocolo com smartphone e iluminação direta apresentou a maior acurácia (97,5%), superando o microscópio digital (91,1%) e o smartphone com iluminação difusa (81,2%).

A hipótese para esse resultado é que a iluminação direta e controlada proporciona contraste mais elevado entre partículas e micro-sombras que ressaltam o relevo superficial das amostras. Isso gera imagens com maior riqueza de detalhes morfológicos, que são cruciais para os algoritmos reconhecerem padrões associados às classes texturais.

Em contrapartida, o método de iluminação difusa, embora elimine reflexos e irregularidades de luz, reduz o contraste e a percepção de profundidade, comprometendo a extração de atributos relevantes, especialmente nas classes Média e Argilosa.

Análise da matriz de confusão

A Figura 1 apresenta a matriz de confusão da melhor combinação do modelo de Regressão Logística com o protocolo de iluminação direta, ilustrando o desempenho preditivo por classe.

Figura 1 – Matriz de confusão (Regressão Logística no Smartphone com Iluminação Direta).

		Predicted				
		Arenoso	Argiloso	Muito Argiloso	Médio	Σ
Actual	Arenoso	100.0 %	0.0 %	0.0 %	6.4 %	91
	Argiloso	0.0 %	98.7 %	2.5 %	0.0 %	78
	Muito Argiloso	0.0 %	1.3 %	97.5 %	0.0 %	78
	Médio	0.0 %	0.0 %	0.0 %	93.6 %	73
Σ		86	77	79	78	320

Fonte: Autoria própria (2025).

Observa-se uma excelente taxa de acerto: 100,0% para a classe Arenoso, 98,7% para Argiloso, 97,5% para Muito Argiloso e 93,6% para Médio, demonstrando uniformidade de desempenho entre as classes predominantes. As pequenas confusões entre as classes Médio e Argiloso refletem a semelhança visual entre as partículas dessas texturas, o que já é conhecido na literatura como um dos maiores desafios em classificações granulométricas automatizadas.

Comparação com métodos convencionais

Os resultados alcançados reforçam que a abordagem por aprendizado de máquina pode substituir parcialmente as análises laboratoriais tradicionais de granulometria, mantendo alta precisão e reduzindo tempo e custo. Enquanto o método da pipeta requer vários dias por amostra e infraestrutura laboratorial, o modelo treinado realiza a classificação em poucos segundos, podendo ser aplicado em campo, via aplicativo móvel.

Além disso, o custo operacional do método proposto é cerca de 90% inferior ao da análise convencional, considerando a eliminação de reagentes e equipamentos especializados.

4. CONCLUSÕES

Conclui-se que a classificação textural do solo por meio de imagens digitais e aprendizado de máquina é uma abordagem viável, precisa e acessível, alcançando 97,5% de acurácia no protocolo de captura via smartphone com iluminação direta demonstrando ser o mais eficaz, representando uma solução de baixo custo e

alta portabilidade, ideal para agilizar a análise de solos em campo.

Os resultados confirmam o potencial dessa tecnologia como ferramenta de apoio à agricultura de precisão, promovendo maior eficiência e sustentabilidade nas práticas de manejo do solo.

Com a expansão do banco de dados para 1.500 amostras, espera-se obter modelos mais robustos e generalizáveis, aprimorando a confiabilidade dos resultados e abrindo caminho para a integração de redes neurais convolucionais (CNNs) em versões futuras, ampliando a capacidade preditiva e a automação do processo de classificação.

REFERÊNCIAS

[1] DONAGEMMA, Guilherme Kangussu; VIANA, João Herbert Moreira; ALMEIDA, Brivaldo Gomes De; RUIZ, Hugo Alberto; KLEIN, Vilson Antônio; DECHEN, Sonia Carmela Falci; FERNANDES, Raphael Bragança Alves. Análise granulométrica. In: TEIXEIRA, Paulo César; DONAGEMMA, Guilherme Kangussu; FONTANA, Aemir; TEIXEIRA, Wenceslau Geraldes (org.). Manual de métodos de análise de solo. 3a ed. Brasília, DF: Embrapa, 2017. p. 573.

Disponível em:

https://www.infoteca.cnptia.embrapa.br/infoteca/handle/do_c/1087262. Acesso em: 30 set. 2022.

AGRADECIMENTOS

Ao CNPq pela bolsa de Iniciação Tecnológica e à Fundação Shunji Nishimura de Tecnologia pelo suporte laboratorial e fornecimento das amostras.

¹Aluno de IT com bolsa CNPq PIBITI - Programa Institucional de Bolsa de Iniciação em Desenvolvimento Tecnológico e Inovação.