

# Reconhecimento de folhas de soja através de descritores de HOG

Murilo Caminotto Barbosa; RA: 2391864

## I. INTRODUÇÃO

O Brasil é um dos grandes produtores de soja do mundo e portanto demanda uma grande variedade de estudos relacionados com esta planta. Em geral, os estudos de soja estão relacionados com sua folha e as possíveis doenças que podem ocorrer em sua superfície. Estas doenças causam lesões severas, o que acarreta na diminuição da absorção de nutrientes, e consequentemente em uma vagem menor e menos densa, causando prejuízos econômicos.

Para a realização de estudos são necessárias diversas amostras da plantas. Estas amostras podem ser desde a folha em si para análises físico-químicas quanto imagens das mesmas para análises computacionais. Em particular para os casos de imagens, para que ocorra uma padronização do posicionamento da folha com relação ao centro da imagem, entre outros quesitos como o foco e luz e sombra, se faz necessário que as folhas sejam removidas das plantas para que sejam levadas até um laboratório em um ambiente controlado. Este processo demanda tempo e recurso, além de ser nocivo para a planta.

Com os avanços tecnológicos nas áreas de processamento digital de imagens e aprendizado de máquina (A.M.), tem se tornado cada vez mais comum criar uma ferramenta para auxílio na coleta de imagens. É sabido que os descritores do tipo Histograma de Gradientes orientados (HOG) são fortes indicadores de formatos e amplamente utilizados para detecção de formas quando aplicados em modelo de A.M..

Caso seja possível treinar um modelo de AM para reconhecimento de imagens de soja em campo, a coleta de imagens para utilização em projetos se tornaria mais simples e rápida, sem a necessidade de coleta física da folha, o que além de prática é menos nocivo à planta.

Abril 13, 2022

## II. OBJETIVOS

O objetivo geral deste trabalho consiste em treinar um modelo de A.M. para o reconhecimento de folhas de soja em imagens de fundo complexo.

Para alcançar o objetivo geral foram propostos os seguintes objetivos específicos:

- Encontrar melhor conjunto de descritores HOG para descrever o problema
- Encontrar melhor conjunto de parâmetros para o modelo classificador utilizado
- Criar imagem ilustrativa contento regiões onde se encontra uma folha de soja

## III. TRABALHOS RELACIONADOS

No trabalho de Thiago T. Santos [3], foram utilizados descritores HOG para diferenciar bagas de café do fundo complexo de uma imagem da planta de café tirada diretamente no campo com iluminação natural. Para tanto, foram classificadas de forma manual 3393 imagens que serviram de treinamento para uma máquina de vetor de suporte. O conjunto ideal de parâmetros para a geração dos descritores HOG foi de 16 células com  $8 \times 8$  pixels e 9 orientações (os descritores HOG serão explicados mais à frente). Através de testes quantitativos foi mostrado uma precisão de 90% na identificação de bagas de café.

## IV. METODOLOGIA EXPERIMENTAL

O foco deste trabalho foi na utilização dos descritores HOG para identificar folhas de soja em imagens de fundo complexo.

### A. Descritores de Histograma de Gradientes Orientados (HOG)

Nos descritores HOG, a distribuição das direções dos gradientes são utilizadas como descritores. A magnitude dos gradientes tende a ser elevada próximo das bordas e cantos, o que os torna úteis para trazer informações sobre o formato dos objetos.

Para calcular os descritores devemos seguir uma série de passos:

- 1) Calcular o gradiente da imagem vertical e horizontal, que pode ser feito através da aplicação do filtro *Sobel* com *kernel* de tamanho 1
- 2) Calcular a magnitude e o ângulo do gradiente seguindo as fórmulas  $g = \sqrt{G_x^2 + G_y^2}$  e  $\Theta = \arctg \frac{G_y}{G_x}$ , onde  $g$  é o gradiente e  $\Theta$  é ângulo.

- 3) Calcular o histograma do gradiente em células de  $N \times N$  para  $K$  orientações. Nesta etapa temos a magnitude dos gradientes e suas direções, portanto podemos calcular a contribuição de cada pixel para o histograma de  $K$  orientações. Isto é feito considerando que o ângulo varia apenas de 0-180. Considerando por exemplo um  $K=9$ , temos ângulos divididos a cada 20°, portanto um ângulo de 10° tem suas magnitudes somadas aos graus 0 e 20 por estar entre eles. Este processo se repete até que tenhamos calculado o histograma para todos os pixels da célula e para todas as células.
- 4) Normalizar o Bloco  $M \times M$  (com  $M \geq N$ ). Se utilizarmos o histograma da forma como calculado acima, variações de luminosidade irão causar distúrbios no resultado. Por este motivo, um processo de normalização entre blocos é aplicado, onde uma quantidade  $X$  de blocos (onde  $X * N = M$ ) é utilizada para calcular a normalização. Então Avança-se  $N$  pixels e se repete o processo, tornando o resultado robusto a mudanças de iluminação.
- 5) Calcular o vetor de gradientes completo. Considerando  $H$  e  $L$  como altura e largura da imagem em questão temos que o total de descritores gerado pelo HOG é  $T = (H/M) * (L/M) * (M * M)$

### B. Banco de imagens

Uma vez que entendemos o que irá descrever a nossa imagem, precisamos de uma base de dados das imagens propriamente ditas. Para a classe positiva, ou seja, de imagens de folhas de soja centralizadas, foi utilizado o banco de imagens público disponível em [1] contendo mais de 65 mil amostras de folhas de soja com diferentes tipos de colorações devido a diferentes níveis de stress, todas com 64x64 pixels.

Para a classe negativa, existe um problema empírico, que consiste no fato de que considerar como algo negativo. Por exemplo, em um caso de imagens tiradas do campo, ervas daninhas e imagens de solo exposto (sem plantas) poderiam ser utilizadas para a classe negativa, enquanto em experimentos em pequena escala, para testes em laboratórios ou universidades (chamados ensaios), teríamos outros objetos para representar a classe negativa, como a textura do chão, mesa, cabos, vasos, caule da planta, outras espécies de plantas, etc. Devido ao tempo disponível para a formulação do trabalho, a aquisição de um banco de imagens próprio para o problema em questão ficou fora de cogitação, portanto um banco de imagens aberto foi utilizado. O banco escolhido foi o *DeepWeeds* disponível em [2], contendo mais de 17 mil amostras de ervas daninhas. As imagens tiveram seu tamanho redefinido para 64x64 para manter o padrão da classe positiva.

### C. Balanceamento das classes

Existe um óbvio desbalanceamento entre as classes negativas e positivas e, para solucionar este problema a seguinte abordagem foi utilizada:

- 1) Para cada imagem da classe negativa foram criadas três (3) outras imagens com rotações em 90° e mais uma (1) imagem invertida.
- 2) Com o total de imagens negativas agora ultrapassando o total de imagens positivas, foram selecionadas de forma aleatória imagens criadas e removidas do conjunto de dados.

Após este processo ambas as classes possuíam a mesma quantidade de imagens, tornando o banco de imagens balanceado.

### D. Modelos de A.M.

Para o aprendizado foram utilizados os classificadores SVM com *kernel RBF* e sigmoide e o classificador *Random Forest*. Para o caso da *Random Forest*, todos os parâmetros foram deixados como padrão.

### E. Criação do modelo

Agora que temos o banco de imagens com mais de 135 mil imagens, entendemos o descritor a ser utilizado para diferenciar cada classe e já definimos os modelos de A.M. a serem testados, resta apenas propor um processo metodológico para a realização dos testes assim como medidas avaliativas para saber se estamos no caminho certo.

Os descritores HOG recebem essencialmente três (3) parâmetros: O número de orientações, o número de pixel por célula e o número de células por bloco. A quantidade de valores que podem ser atribuídos a cada um destes parâmetros varia de acordo com a resolução da imagem, porém, de toda forma estes parâmetros possuem complexidade combinatória o que torna impraticável testar todas as possibilidades. Portanto precisamos dar um “chute inicial” de forma orientada pela literatura.

No trabalho citado acima, a combinação de 9 orientações com com 8 pixel por células e um total de 2 células por blocos (16 pixels) foi aplicada, portanto parece razoável começarmos por aí. Para cada folha do banco de imagens, os descritores HOG foram extraídos e colocados em um arquivo **CSV** juntamente com sua classe sendo o valor 1 para classe positiva (folha de soja) e 0 para classe negativa.

O modelo é então treinado utilizando validação cruzada estratificada com 33% das amostras para teste, treinando portanto com 44059 amostras de cada classe resultando em um arquivo salvo no formato **joblib** para utilização futura em casos reais

Para avaliar o desempenho do classificador, foram utilizadas as medidas de **Acurácia** e **F1-score**. Na seção de Resultados e discussões será apresentada uma tabela contendo todas as combinações de descritores utilizados assim como seus desempenhos, mas por hora vamos nos ater ao procedimento metodológico e no raciocínio para a criação de novos modelos.



Fig. 2. Imagem de campo retirada da internet.



Fig. 3. Imagem de ensaio de soja.

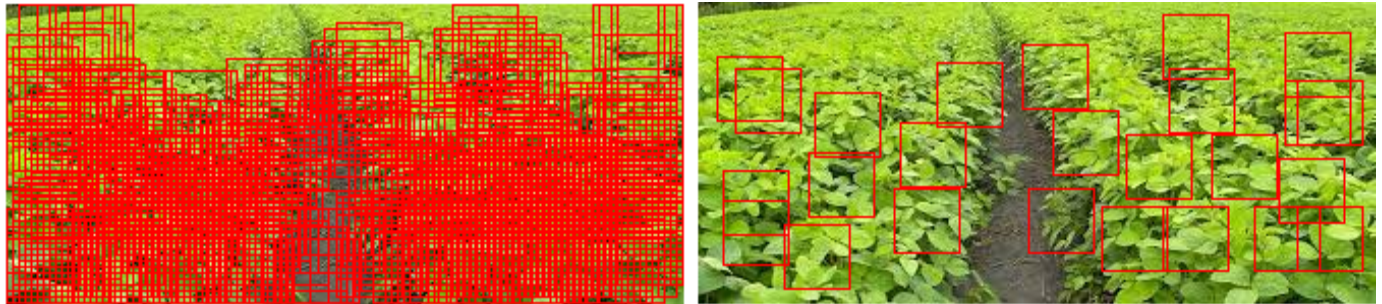


Fig. 4. Resultado do modelo de A.M. aplicado na imagem de campo; A esquerda não foi aplicada a limpeza de sobreposições; A direita foi aplicada a limpeza de sobreposições

Para realizar a classificação em imagens reais, dada uma imagem de tamanho  $H \times L$ , ela era recortada em quadrados de 32, 64, 128 e 364 pixels, sendo estas as medidas utilizadas para encontrar possíveis regiões contendo folhas. Cada quadrado era retirado a uma distância de 10% de seu tamanho, ou seja, um quadrado de  $32 \times 32$  retirado na posição (0,0) teria seus próximos vizinhos nas posições (3,0) e (0,3), aumentando as chances de encontrar uma região com classe positiva. Caso uma classe positiva seja encontrada, um quadrado vermelho é pintado. É esperado que acontecesse grande quantidade de sobreposição já que quadrados vizinhos estão próximos um do outro, portanto uma forma de reduzir sobreposições foi proposta e está explicada na figura 1

A imagem de estudo de caso foi retirada da internet com acesso livre e além da imagem de campo, 15 imagens de ensaios agrícolas foram gentilmente cedidas pela doutoranda em Bioinformática pela UTFPR de Cornélio Bruna Avelino Barbosa, que realiza pesquisas na Embrapa com sede em Londrina. A imagem de campo e um exemplo de imagem de ensaio são exibidos nas figuras 2 IV-E respectivamente.

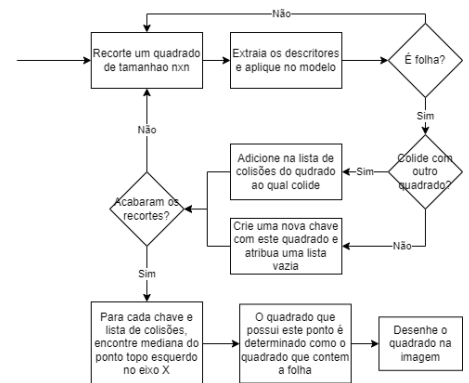


Fig. 1. Fluxo para remoção de sobreposição de quadrados

Após os resultados excelentes do primeiro modelo testado (9,8,16) com Acurácia e F1-scores acima de 99%, era de se esperar que o modelo obtido fosse capaz de encontrar com facilidade imagens de folha de soja dada uma imagem de campo real, porém não foi bem isso o que ocorreu, conforme ilustrado pela figura 4

Embora o modelo tenha aprendido a diferenciar entre as amostras positivas e negativas com maestria, não foi capaz de generalizar para uma nova imagem nunca vista antes. Isto pode ser devido a diversos fatores, dentre eles: ruídos, iluminação, contrastes, cores entre outros. Para tentar resolver esse problema diversas abordagens foram testadas:

- 1) Diminuir a quantidade de orientações na tentativa de mitigar o efeito de diferença de iluminação.
- 2) Equalizar histograma para normaliza a iluminação das imagens
- 3) Binarização para melhora o contorno da imagem
- 4) Transferência de contraste descrito em [4] definido por:
  - a) Cálculo da média e desvio padrão de todas as imagens do banco de imagens. Serão chamados de *mediaGlobal* e *stdGlobal*.
  - b) Para cada imagem, será realizada uma transformação de padronização de contraste que consiste em:
    - i) Realizar a equalização do histograma

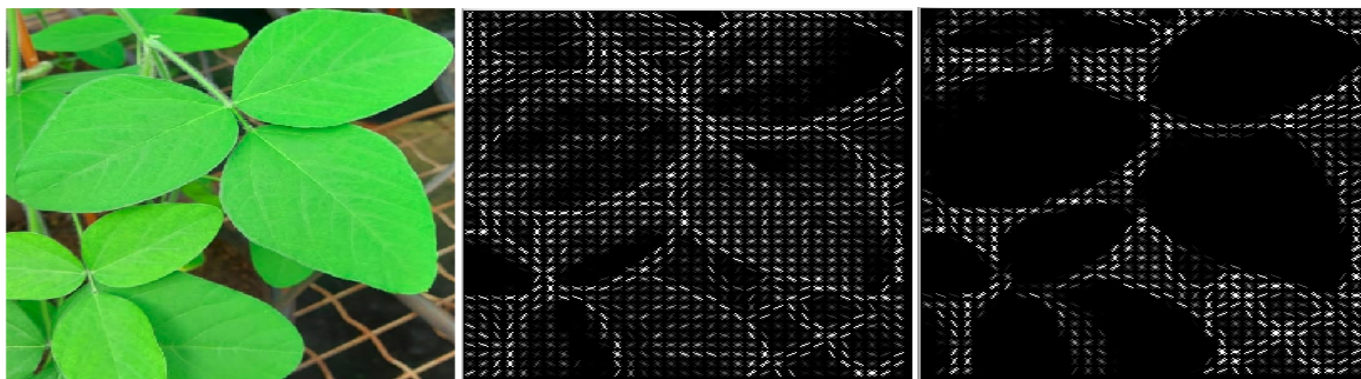


Fig. 5. Esquerda: Imagem original; Centro: Descritores HOG da imagem equalizada; Direita: Descritores HOG da imagem com transferência de contraste;

- ii) Criação de uma tabela de *lookup* a nível de pixel a partir da função:  $\text{ novoValor} = A * \text{valor} + B$ , onde  $A = \frac{\text{std}(\text{imagequalizada})}{\text{stdGlobal}}$  e  $B = \text{mean}(\text{imagequalizada}) - A * \text{mediaGlobal}$

A figura 5 representa alguns dos processos descritos acima, e para cada processo, uma ou mais tabelas de descritores **HOGs** foram criadas e utilizadas para treinar os diferentes modelos. Os resultados são mostrados a seguir

## V. RESULTADOS E DISCUSSÕES

A tabela 1 a seguir contém os resultados de cada planilha criada, onde o nome da planilha representa os parâmetros utilizados mais o pré processamento quando aplicado.

Planilha	SVM RBF		SVM Sigmoid		Random Forest	
	F1-Score	Acurácia	F1-Score	Acurácia	F1-Score	Acurácia
database-4-8-2	0.9998	0.9998	0.9997	0.9997	0.9999	0.9999
database-4-9-2	0.9997	0.9997	0.9997	0.9996	0.9999	0.9999
database-4-9-7	0.9998	0.9998	0.9996	0.9997	0.9998	0.9998
database-9-8-8	0.9998	0.9998	0.9997	0.9997	0.9999	0.9999
database-9-8-2-transformed	0.9995	0.9995	0.9994	0.9994	0.9996	0.9996
database-4-8-1-bin-transformed	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
database-4-4-1-hist-only	0.9995	0.9995	0.9994	0.9994	0.9998	0.9998
database-4-4-1-bin-transformed	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

O motivo mais provável para o resultado obtido é que o conjunto de amostras negativas utilizadas foram imagens de ervas daninhas, com o intuito de diferenciar outros tipos de vegetação, assim como galhos e solo exposto. Porém as imagens utilizadas para teste não possuem estes tipos de objetos, o que provavelmente confundiu o classificador.

No projeto proposto em [3], as imagens utilizadas para criar o classificador foram retiradas no formato desejado e específico para a aplicação. Como as imagens utilizadas neste projeto foram retiradas de bancos diferentes com finalidades diferentes, o resultado acabou por não ser capaz de classificar corretamente as imagens de teste.

## VI. CONCLUSÃO

Embora tenha sido possível alcançar métricas excelentes para a classificação entre positivo e negativo, a classe negativa não representou corretamente o problema real. Para um desempenho aprimorado, acredita-se que exemplos retirados de imagens semelhantes às que serão analisadas devam ser manualmente catalogadas de forma que as classes negativas sejam melhor representadas. Além disso, a etapa de transferência de contraste resultou em uma melhor diferenciação das classes quando na imagem real. Para trabalhos futuros, talvez a melhor alternativa seja coletar mais amostras de imagens específicas para uma determinada aplicação, como por exemplo coleta de dados em ensaios de soja, tornando o modelo mais específico para o trabalho em questão.

## REFERENCES

- [1] GHOSAL, Sambuddha et al. An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences*, v. 115, n. 18, p. 4613-4618, 2018.
- [2] OLSEN, Alex et al. DeepWeeds: A multiclass weed species image dataset for deep learning. *Scientific reports*, v. 9, n. 1, p. 1-12, 2019.
- [3] SANTOS, Thiago T. Detecção automática de bagas de café em imagens de campo. In: *Embrapa Informática Agropecuária-Artigo em anais de congresso (ALICE)*. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 10., 2015, Ponta Grossa. *Uso de VANTs e sensores para avanços no agronegócio: anais*. Ponta Grossa: Universidade Estadual de Ponta Grossa, 2015., 2015.
- [4] ROSEBROCK, Adrian. Super fast color transfer between images. Disponível em: <https://pyimagesearch.com/2014/06/30/super-fast-color-transfer-images/>; Acessado em 13.04.2022.