

Reconhecimento de folhas de soja através de descritores de HOG

Murilo Caminotto Barbosa

Universidade Tecnológica Federal do Paraná

Cornélio Procópio, 2022

- 1 Introdução
- 2 Objetivos
- 3 Trabalhos Relacionados
- 4 Metodologia
- 5 Resultados e discussões
- 6 Conclusão

- 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.
- Para a realização de estudos são necessárias diversas amostras das plantas.
- Este processo demanda tempo e recursos, além de ser nocivo a planta.

- Com os avanços nas áreas de processamento digital de imagens e aprendizado de máquina (A.M.), tem se tornado cada vez mais popular a criação de uma ferramenta para auxílio na coleta de imagens.
- E sabido que os descritores do tipo Histograma de Gradientes Orientados (**HOG**) são fortes indicadores de formatos
- 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

Objetivos

Geral

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.

Específicos

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

Exemplo de resultado esperado

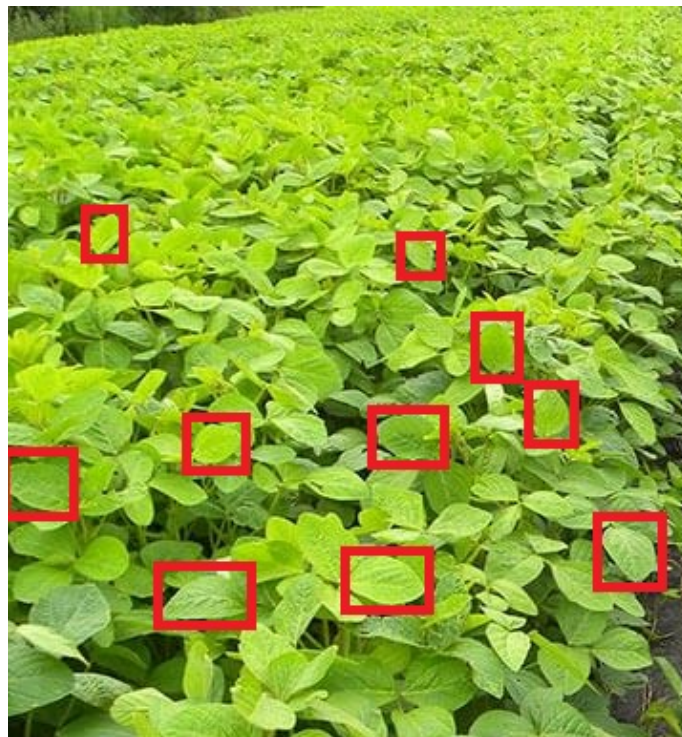


Figura 1: Exemplo ilustrativo do resultado esperado. Apenas as folhas que se destaquem mais são selecionadas e podem ser extraídas do total da imagem.

- No trabalho de Thiago T. Santos [3], foram utilizados descritores **HOG** para diferenciar bagas de café do fundo complexo de imagens da planta tiradas 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 8×8 pixels e 9 orientações
- Através de testes quantitativos foi mostrado uma precisão de 90% na identificação de bagas de café

Exemplo de imagem retirada do trabalho

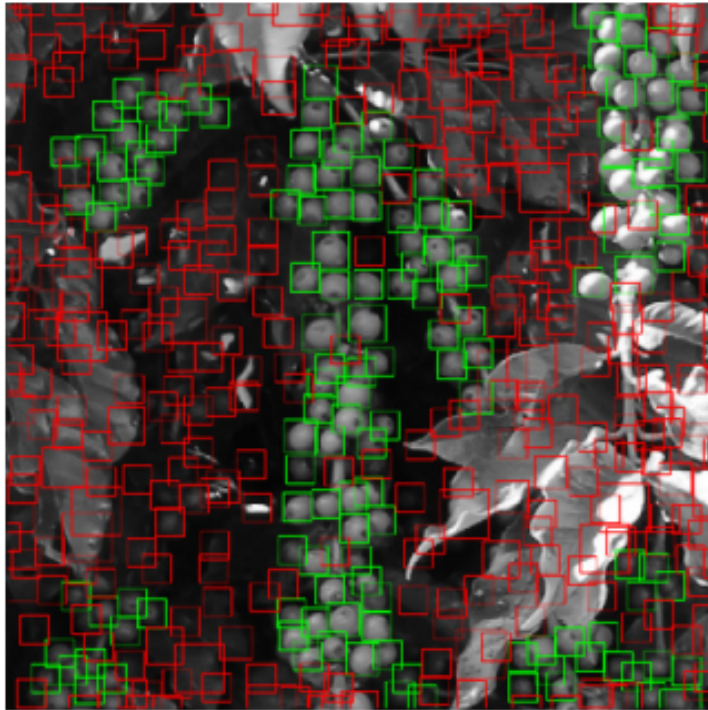


Figura 2: Exemplo de imagem resultado classificando bagas de café em imagem de campo com iluminação natural

Descritores de Histograma de gradientes orientados (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.

A seguir serão descritores os passo que devem ser tomados para calcular estes descritores

- Calcular o gradiente da imagem vertical e horizontal, que pode ser feito através da aplicação do filtro *Sobel* com *kernel* de tamanho 1.



Figura 3: Filtro Sobel

- Calcular a magnitude e o ângulo do gradiente seguindo as fórmulas: $g = \sqrt{Gx^2 + Gy^2}$ e $\Theta = \arctg \frac{Gy}{Gx}$, onde g é o gradiente e Θ é ângulo.

- 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.

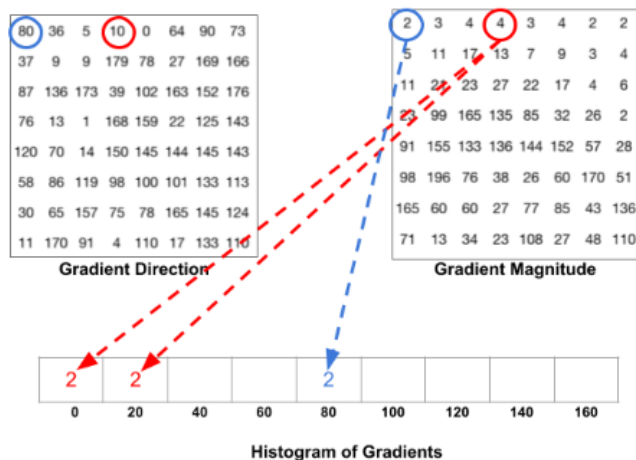


Figura 4: Processo de calculo do histograma de gradientes

- Normalizar o Bloco $M \times M$ (com $M \geq N$). 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.

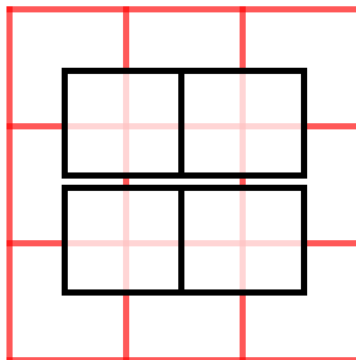


Figura 5: Ilustração do processo de normalização

- Calcular o vetor de gradientes completo. Considerando H e L como altura de largura da imagem em questão temos que o total de descritores gerado pelo HOG é $T = (H/M) * (L/M) * (M * M)$

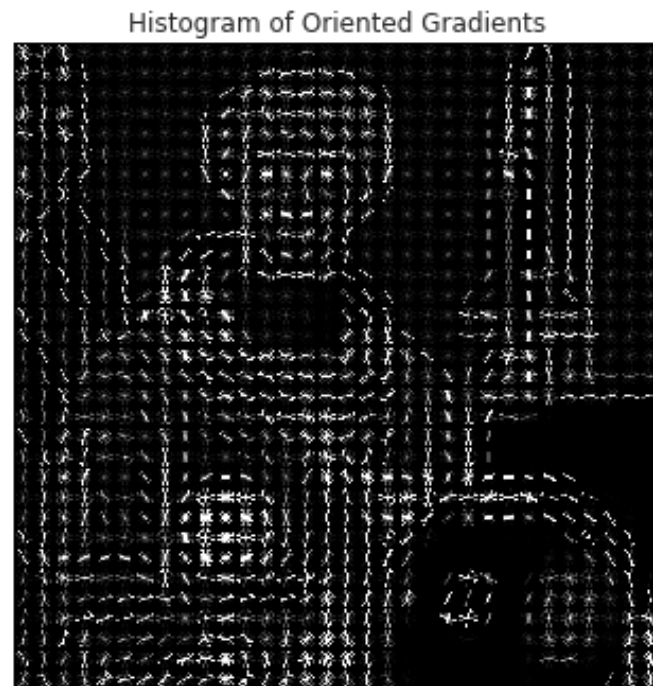


Figura 6: Exemplo de resultado de imagem após o calculo dos histograma orientado a gradiente.

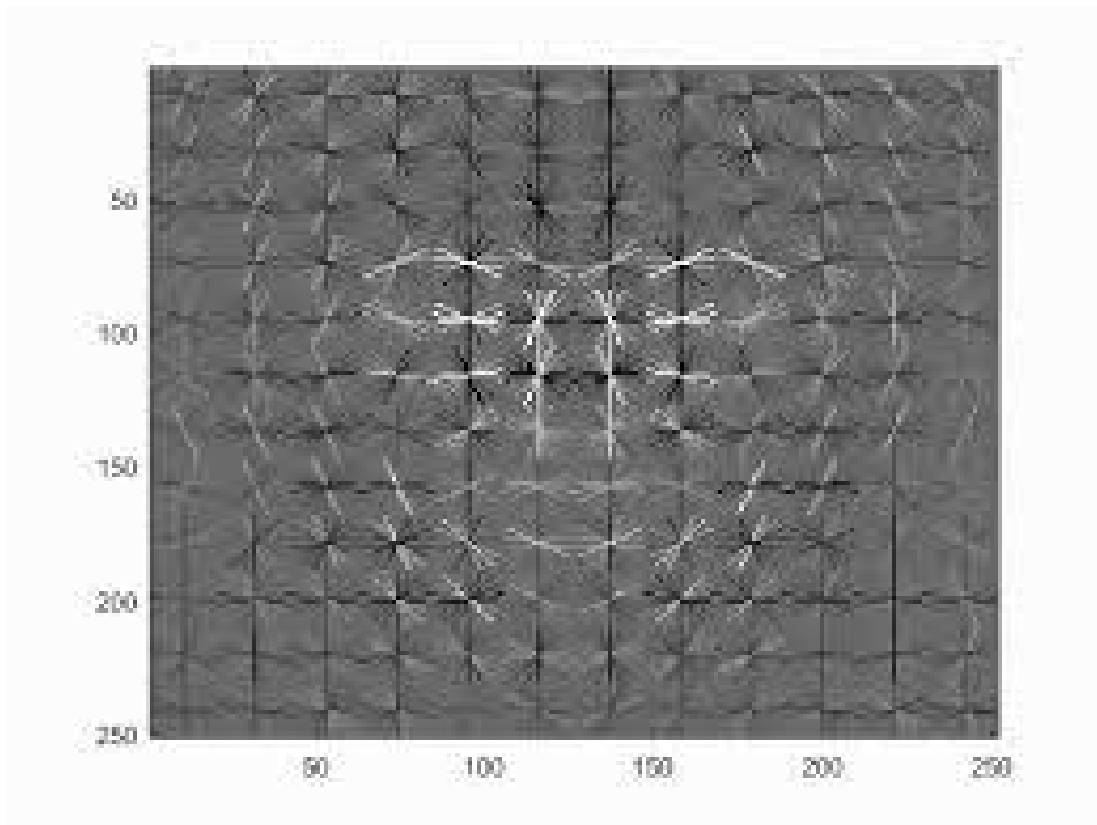
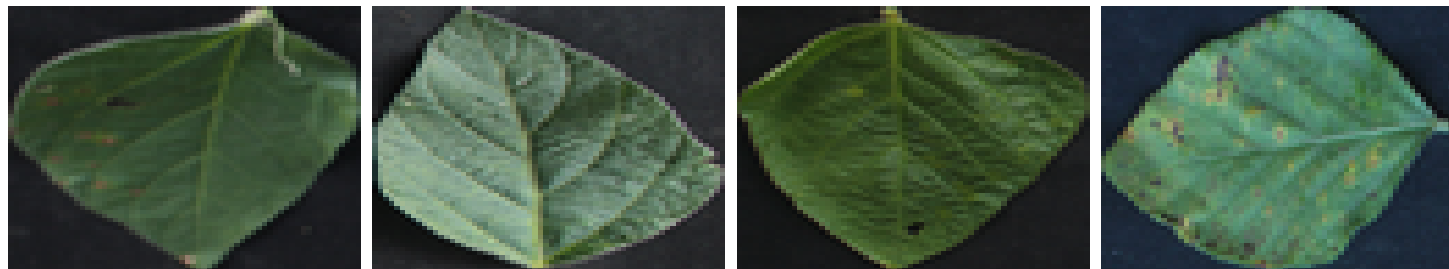


Figura 7: Descritores HOG combinados para detecção de rostos em imagens.

Banco de imagens

Classe Positiva (Folha)

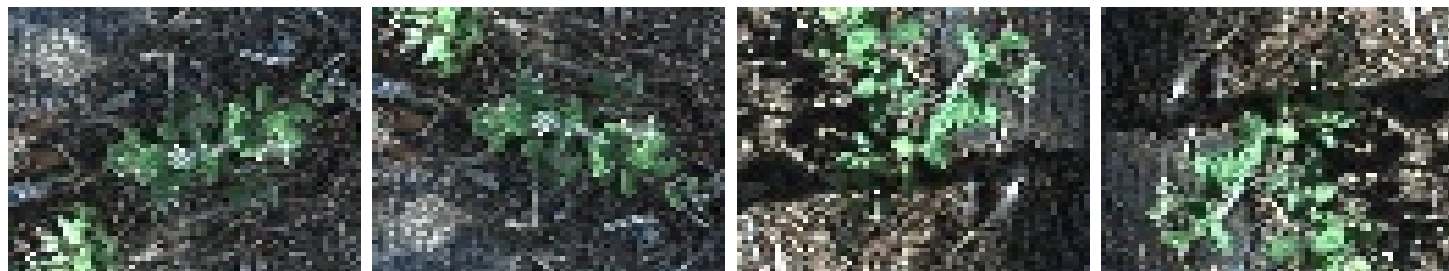
- Para a classe positiva, ou seja, de imagens de folhas de soja centralizadas, foi utilizado o banco de imagens público disponível em [1].
- Este banco de imagens contém mais de 65 mil amostras de folhas de soja com diferentes tipos de colorações devido a diferentes níveis de stress, todas com 64x64 pixels.



Banco de imagens

Classe Negativa

- Para a classe negativa, existe um problema empírico, que consiste no fato de que considerar como algo negativo.
- Devido ao tempo disponível para a formulação do trabalho, a aquisição de uma banco de imagens próprio para o problema em questão ficou fora de cogitação.
- O banco escolhido foi o *DeepWeeds* disponível em [2], contendo mais de 17 mil amostras de ervas daninhas.
- Todas as imagens foram redimensionadas para 64x64.



Balanceamento das classes

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

- ❶ Para cada imagem da classe negativa foram criadas três (3) outras imagens com rotações em 90° e mais uma (1) imagem invertida.
- ❷ 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

Apos este processo ambas as classes possuem a mesma quantidade de imagens, tornando o banco de imagens balanceado

Para o aprendizado foram utilizados os classificadores:

- SVM Kernel RBF
- SVM Kernel Sigmoid
- *Random Forest*

Para o caso da *Random Forest*, todos os parâmetros foram deixados como padrão.

Criação do arquivo de amostras

- Para criação do arquivo de descritores precisamos decidir quais parâmetros utilizar para o **HOG**
- A complexidade de escolha é combinatória
- Foi decidido como ‘chute inicial’ utilizar as configurações sugeridas em [3]: 9 orientações, 8 pixels por célula e 2 celular por bloco (16 pixels)
- Para cada uma das mais de 135 mil folhas, os descritores foram extraídos e escrito num arquivo **.csv**
- A classe positiva foi representada pelo valor **1** e a negativa pelo valor **0**

Treinamento do modelo

- O modelo foi treinado através de validação cruzada estratificada com 33% das amostras para teste, resultado em 44059 amostras de cada classe para treino.
- Após o modelo ser treinado, ele é salvo num arquivo de formato **joblib**
- Para avaliar o desempenho do modelo foram adotadas as medidas de **Acurácia** e **F1-Score**

Como encontrar regiões com folhas?

- 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
- Cada quadrado era retirado a uma distância de 10% de seu tamanho, ou seja, um quadrado de 32×32 retirado na posição (0,0) teria seus próximos vizinhos nas posições (3,0) e (0,3)
- Os descritores são então retirados para cada quadrado e aplicadas no modelo
- Caso uma classe positiva seja encontrada, um quadrado vermelho é pintado

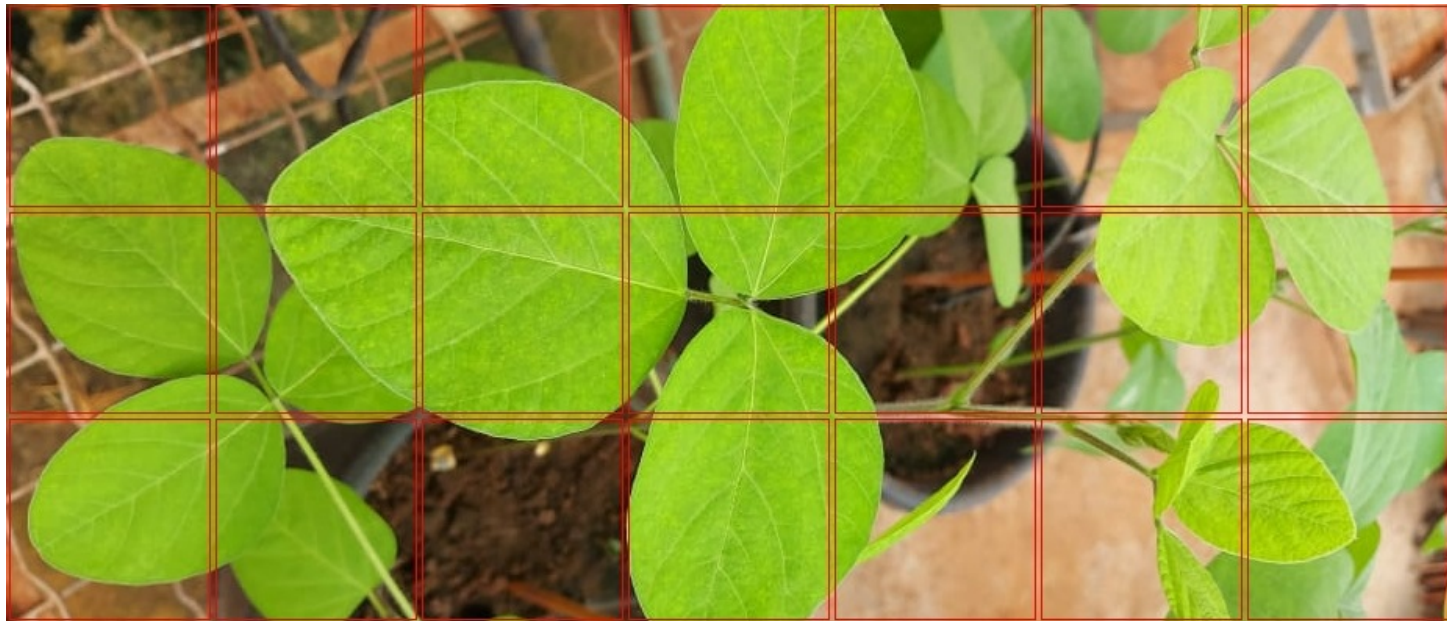


Figura 8: Ilustração de quadrados recortados em uma imagem

Remoção de sobreposições

- 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 a seguir

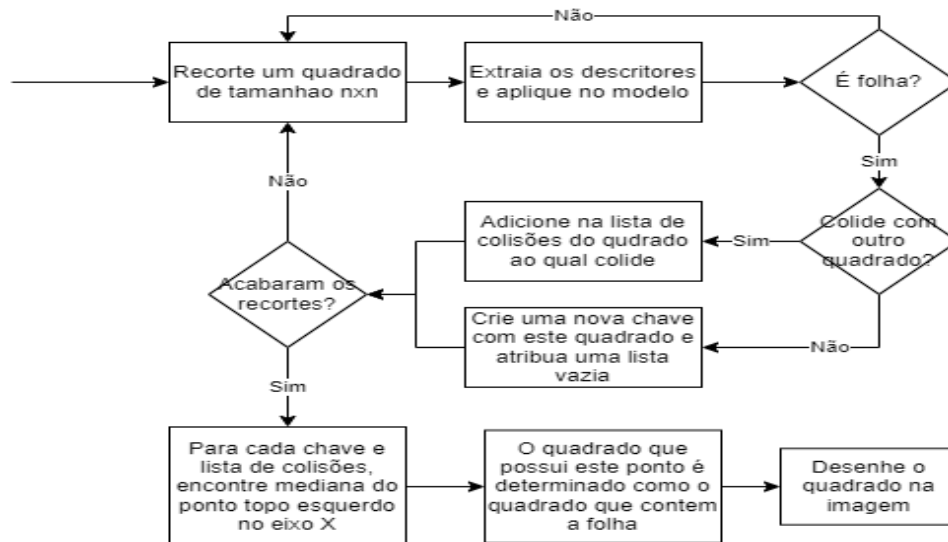


Figura 9: Fluxo de remoção de quadrados com sobreposições

Imagens de teste

- A imagem de estudo de caso foi retirada da internet com acesso livre
- Além da imagem de campo, 15 imagens de ensaios agrícolas foram gentilmente cedidas por uma doutoranda em Bioinformática pela UTFPR, que realiza pesquisas na Embrapa relacionadas a soja.



Figura 10: Imagem de campo retirada da internet.



Figura 11: Imagem de ensaio de soja.

Exemplo de Resultado

- Após os resultados do primeiro modelo testado (9,8,16) com acurácia e F1-scores acima de 99%, era esperado que o modelo obtido fosse capaz de encontrar as imagens de folha de soja dada uma imagem de campo real.
- Porém não foi isso o que aconteceu conforme ilustrado abaixo

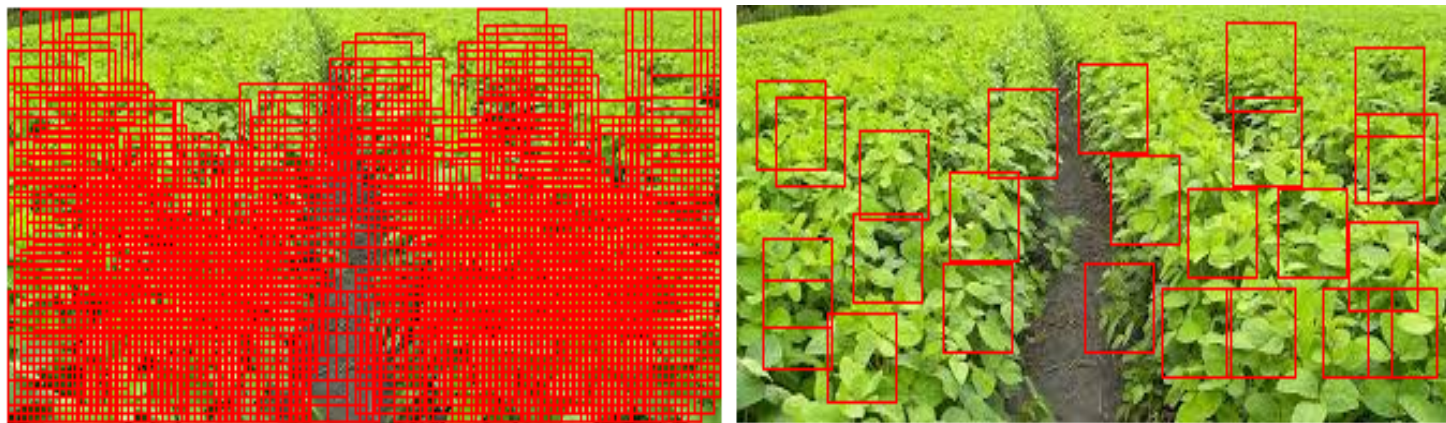


Figura 12: Resultado do modelo de A.M. aplicado na imagem de campo; A esquerda não foi aplicada a limpeza de sobreposições; A direita foi aplicada a limpeza de sobreposições

Possíveis motivos pelo qual falhou

- Embora o modelo tenha aprendido a diferenciar entre as amostras positivas e negativas, não foi capaz de generalizar para uma nova imagem
- Isto pode ser devido a diversos fatores, dentre eles:
 - 1 Ruídos
 - 2 Iluminação
 - 3 Contrastes
 - 4 Cores
 - 5 Entre outros

Para tentar resolver esse problema diversos pré-processamentos foram testados:

Pré-processamentos aplicados

- ❶ Diminuir a quantidade de orientações na tentativa de mitigar o efeito de diferença de iluminação.
- ❷ Equalizar histograma para normalizar a iluminação das imagens
- ❸ Binarização para melhora o contorno da imagem
- ❹ Transferência de contraste descrito em [4] definido por:
 - ❶ Cálculo da média e desvio padrão de todas as imagens do banco de imagens. Serão chamados de *mediaGlobal* e *stdGlobal*.
 - ❷ Para cada imagem, será realizada uma transformação de padronização de contraste que consiste em:
 - ❶ Realizar a equalização do histograma
 - ❷ Criação de uma tabela de *lookup* a nível de pixel a partir da função:
$$\text{novoValor} = A * \text{valor} + B, \text{ onde } A = \frac{\text{std}(\text{imgequalizada})}{\text{stdGlobal}} \text{ e } B = \text{mean}(\text{imgequalizada}) - A * \text{mediaGlobal}$$

Exemplo de imagens pós processamento

- A figura a seguir ilustra alguns dos pré-processamentos aplicados

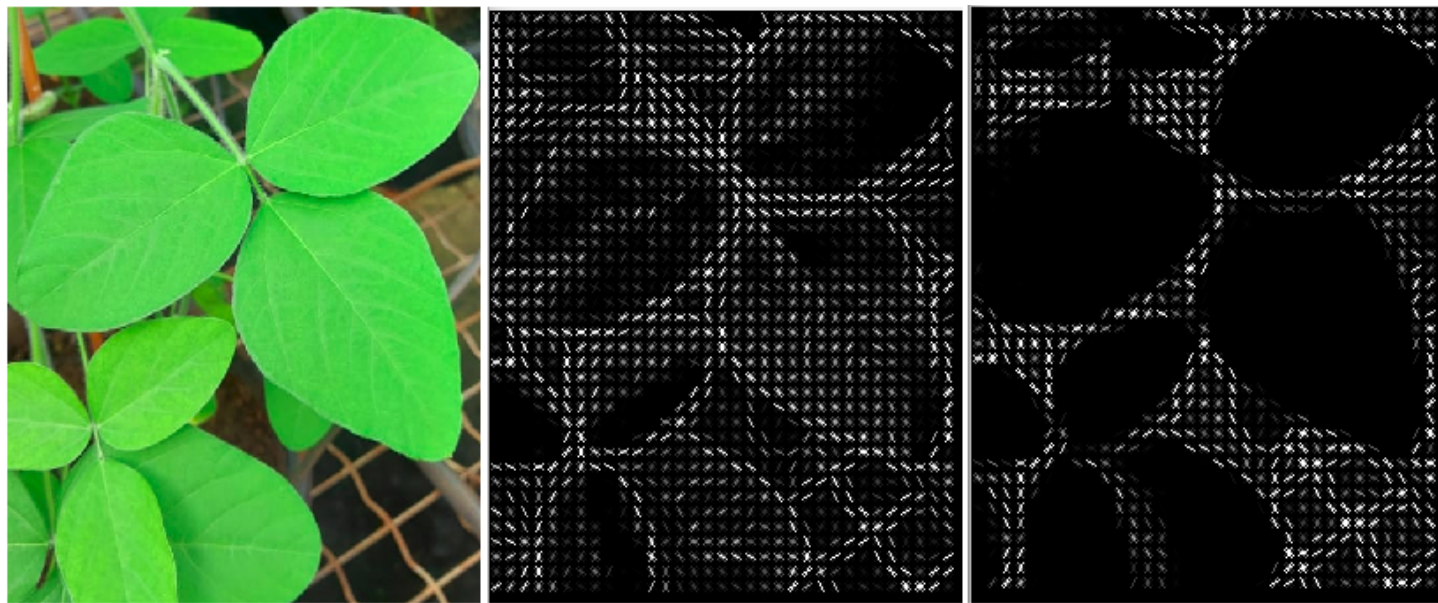


Figura 13: Esquerda: Imagem original; Centro: Descritores HOG da imagem equalizada; Direita: Descritores HOG da imagem com transferência de contraste;

Criação de novos modelos

- Para cada pré-processamento aplicado uma ou mais tabelas de descritores foi criada
- Para cada tabela, todos os modelos foram retreinados e testados
- A seguir veremos os resultados obtidos por cada tipo de tabela

Resultados e discussões





A Tabela 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 foi 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.
- 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

- Embora tenha sido possível alcançar métricas excelentes para a classificação entre positivo e negativo, a classe negativa não representou corretamente o problema real.
- A etapa de transferência de contraste resultou em uma melhor diferenciação das classes quando na imagem real
- Para trabalhos futuros, a provável melhor alternativa seja coletar mais amostras de imagens específicas para uma determinada aplicação.

Referências

-  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.
-  OLSEN, Alex et al. DeepWeeds: A multiclass weed species image dataset for deep learning. Scientific reports, v. 9, n. 1, p. 1-12, 2019.
-  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 AGROINFORMATICA, 10., 2015, Ponta Grossa. Uso de VANTs e sensores para avanços no agronegócio: anais. Ponta Grossa: Universidade Estadual de Ponta Grossa, 2015., 2015.
-  ROSEBROCK, Adrian. Super fast color transfer between images. Dispon ´ível em: <https://pyimagesearch.com/2014/06/30/super-fast-color-transfer-images/> ; Acessado em 13.04.2022

Obrigado!