REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218

# RECONHECIMENTO DE DOENÇAS EM FOLHAS DE MACIEIRA UTILIZANDO *BAG OF FEATURES* E SVM COM VARIAÇÃO DO VOCABULÁRIO VISUAL

## *RECOGNITION OF DISEASES IN APPLE TREE LEAVES USING BAG OF FEATURES AND SVM WITH VISUAL VOCABULARY VARIATION*

## *RECONOCIMIENTO DE ENFERMEDADES EN HOJAS DE MANZANO UTILIZANDO BAG OF FEATURES Y SVM CON VARIACIÓN DEL VOCABULARIO VISUAL*

Leandro Aureliano da Silva[1], Eduardo Silva Vasconcelos[2], Adriano Dawison de Lima[3], Edilberto Pereira Teixeira[4]

e727212
https://doi.org/10.47820/recima21.v7i2.7212

PUBLICADO: 02/2026

**RESUMO**

O reconhecimento automático de doenças em plantas tem se consolidado como uma ferramenta fundamental para a agricultura de precisão, especialmente em culturas de elevado valor comercial, como a macieira (*Malus domestica*). Neste trabalho é proposta uma abordagem de classificação baseada no modelo *Bag of Features* (BoF) aliado a Máquinas de Vetores de Suporte (SVM), aplicada ao *dataset Apple Tree Leaf*, que contém quatro classes patológicas: *Alternaria leaf spot*, *Brown spot*, *Gray spot* e *Rust*. O estudo diferencia-se pela investigação sistemática do impacto do tamanho do vocabulário visual na etapa de quantização das características, avaliando cinco configurações de BoF (100, 300, 700, 1000 e 3000 palavras). Os experimentos conduzidos no MATLAB R2023b revelam que vocabulários intermediários, em especial entre 700 e 1000 termos, proporcionam o melhor compromisso entre discriminação e capacidade de generalização, resultando em elevadas taxas de acurácia para a maioria das classes. Apesar disso, a classe *Gray spot* demonstrou maior dificuldade de reconhecimento, reforçando a necessidade de estratégias complementares para sua distinção. Os achados confirmam a eficiência de métodos clássicos de visão computacional em bases de tamanho moderado e oferecem subsídios relevantes para o desenvolvimento de sistemas de diagnóstico foliar aplicados ao contexto agrícola.

**PALAVRAS-CHAVE**: Visão computacional. *Bag of Features*. SVM. Doenças em plantas.

***ABSTRACT***

*Automatic recognition of plant diseases has become a fundamental tool for precision agriculture, especially in crops of high commercial value, such as apple trees (Malus domestica). In this work, a classification approach based on the Bag of Features (BoF) model combined with Support Vector Machines (SVM) is proposed and applied to the Apple Tree Leaf dataset, which contains four pathological classes: Alternaria leaf spot, Brown spot, Gray spot, and Rust. The study is distinguished by a systematic investigation of the impact of visual vocabulary size during the feature quantization stage, evaluating five BoF configurations (100, 300, 700, 1000, and 3000 words). The experiments, conducted in MATLAB R2023b, reveal that intermediate vocabularies—particularly between 700 and 1000 terms—provide the best trade-off between discriminative power and generalization capability, resulting in high accuracy rates for most classes. Nevertheless, the Gray spot class exhibited greater difficulty in recognition, reinforcing the need for complementary strategies for its discrimination. The findings confirm the effectiveness of classical computer vision*

[1]Doutor em Engenharia Elétrica, Universidade de Uberaba, Uberaba-MG, Brasil.
[2]Doutor em Engenharia Elétrica, Instituto Federal Goiano, Cristalina-GO, Brasil.
[3]Doutor em Agronomia, Universidade de Uberaba, Uberaba-MG, Brasil.
[4]Doutor em Educação, Universidade de Uberaba, Uberaba-MG, Brasil.


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

*methods on moderately sized datasets and provide relevant insights for the development of leaf diagnosis systems applied to agricultural contexts.*

***KEYWORDS:*** *Computer vision. Bag of Features. SVM (Support Vector Machine). Plant diseases*

***RESUMEN***
*El reconocimiento automático de enfermedades en plantas se ha consolidado como una herramienta fundamental para la agricultura de precisión, especialmente en cultivos de alto valor comercial, como el manzano (Malus domestica). En este trabajo se propone un enfoque de clasificación basado en el modelo Bag of Features (BoF) combinado con Máquinas de Vectores de Soporte (SVM), aplicado al conjunto de datos Apple Tree Leaf, que contiene cuatro clases patológicas: Alternaria leaf spot, Brown spot, Gray spot y Rust. El estudio se diferencia por la investigación sistemática del impacto del tamaño del vocabulario visual en la etapa de cuantización de características, evaluando cinco configuraciones de BoF (100, 300, 700, 1000 y 3000 palabras). Los experimentos, realizados en MATLAB R2023b, muestran que los vocabularios intermedios—especialmente entre 700 y 1000 términos—proporcionan el mejor equilibrio entre capacidad discriminativa y generalización, lo que resulta en altas tasas de exactitud para la mayoría de las clases. No obstante, la clase Gray spot presentó mayor dificultad de reconocimiento, lo que refuerza la necesidad de estrategias complementarias para su diferenciación. Los hallazgos confirman la eficacia de los métodos clásicos de visión por computadora en conjuntos de datos de tamaño moderado y aportan contribuciones relevantes para el desarrollo de sistemas de diagnóstico foliar aplicados al contexto agrícola.*

***PALABRAS CLAVE****: Visión por computadora. Bag of Features. SVM (Máquina de Vectores de Soporte). Enfermedades de las plantas.*

## INTRODUÇÃO

A agricultura de precisão tem incorporado, de forma crescente, técnicas de visão computacional e aprendizado de máquina como suporte ao monitoramento automatizado de culturas, à detecção precoce de doenças e à otimização da produtividade agrícola. No contexto da cultura da macieira (*Malus domestica*), enfermidades foliares como *Alternaria leaf spot*, *Brown spot*, *Gray spot* e *Rust* representam importantes fatores de redução de rendimento e qualidade dos frutos, tornando essencial o desenvolvimento de ferramentas tecnológicas capazes de apoiar o diagnóstico fitossanitário de forma rápida, objetiva e acessível (Mohanty; Hughes; Salathé, 2016; Zhong; Zhao, 2020).

Nas últimas décadas, modelos computacionais para reconhecimento de padrões evoluíram significativamente, abrangendo desde descritores manuais, como SIFT e SURF, até arquiteturas baseadas em redes neurais profundas (Lecun; Bengio; Hinton, 2015). Embora as redes convolucionais apresentem elevado desempenho quando aplicadas a grandes bases de dados, sua adoção enfrenta limitações relacionadas ao custo computacional e à necessidade de conjuntos extensos de dados rotulados (Ferentinos, 2018). Nesse cenário, métodos clássicos como o modelo *Bag of Features* (BoF) associado às Máquinas de Vetores de Suporte (SVM) permanecem relevantes, sobretudo em bases moderadas e com predominância de padrões texturais, como é o caso de imagens de folhas (Pawara *et al.*, 2017; Pires *et al.,* 2016).


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Entretanto, conforme apontam Yang *et al.,* (2007) e Huang *et al.,* (2014), um dos fatores críticos para o desempenho de modelos baseados em *Bag of Features/Bag of Visual Words* está associado ao tamanho do vocabulário visual empregado na quantização das características. Vocabulários reduzidos tendem a limitar a expressividade da representação e comprometer a discriminação entre classes, enquanto vocabulários excessivamente grandes podem aumentar a sensibilidade a ruídos, elevar a dimensionalidade e favorecer o sobreajuste, reduzindo a capacidade de generalização. Apesar desse consenso sobre o *trade-off*, persiste uma lacuna na literatura quanto à definição de configurações ótimas de vocabulário visual em aplicações agrícolas específicas, como o reconhecimento de doenças foliares em macieiras.

Diante desse contexto, o problema que orienta este estudo consiste em compreender de que maneira a variação do tamanho do vocabulário visual influencia o desempenho de sistemas de reconhecimento automático de doenças em folhas de macieira baseados na combinação BoF e SVM, bem como identificar quais configurações oferecem melhor equilíbrio entre capacidade discriminativa e generalização.

A realização deste trabalho justifica-se tanto pela sua relevância científica quanto aplicada, uma vez que contribui para o aprofundamento da compreensão sobre o comportamento de métodos clássicos de visão computacional em contextos agrícolas reais. Além disso, os resultados obtidos oferecem subsídios técnicos para o desenvolvimento de sistemas de diagnóstico mais acessíveis, interpretáveis e viáveis em ambientes com restrições de dados ou infraestrutura computacional, conforme apontado por Barbedo (2016) e Abade, Ferreira e Vidal (2021).

Dessa forma, o objetivo geral deste trabalho é avaliar a eficácia do modelo *Bag of Features* associado ao classificador SVM na identificação automática de doenças em folhas de macieira, considerando a influência do tamanho do vocabulário visual no desempenho da classificação.

Como objetivos específicos, busca-se: (i) aplicar a abordagem BoF+SVM ao *dataset Apple Tree Leaf*, composto por diferentes classes de doenças foliares; (ii) investigar o comportamento do desempenho do classificador frente à variação do vocabulário visual em diferentes configurações; (iii) analisar as taxas de acerto por classe, identificando padrões de confusão entre as doenças; e (iv) discutir as potencialidades e limitações do método proposto no contexto de aplicações voltadas à agricultura de precisão.

## REFERENCIAL TEÓRICO

### *Bag of Features* (BoF)

O método *Bag of Features*, introduzido por Csurka *et al.,* (2004), adapta o *Bag of Words* da linguística para imagens, representando cada imagem como um histograma de “palavras visuais”. Ele segue quatro etapas fundamentais:


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

- Detecção de pontos de interesse: normalmente com SIFT, SURF ou detectores automáticos. Neste trabalho foi utilizado o SURF.
- Extração de descritores locais: cada região da imagem produz um vetor numérico de características.
- Agrupamento *(k-means):* produz-se um conjunto de k “termos visuais”, formando o vocabulário.
- Codificação: cada imagem é quantizada segundo o vocabulário, gerando um histograma de frequências.

O tamanho do vocabulário visual condiciona o poder discriminativo da representação: vocabulários reduzidos tendem a gerar descrições menos informativas, enquanto vocabulários muito extensos podem aumentar a separabilidade, mas também elevar a sensibilidade a ruído e favorecer sobreajuste, reduzindo a capacidade de generalização (Huang *et al.,* 2014).

Segundo Nowak, Jurie e Triggs (2006), representações do tipo *Bag of Features* tendem a ser particularmente eficazes quando as classes apresentam regularidades estatísticas locais (frequentemente associadas a padrões de textura), pois a distribuição de descritores quantizados captura diferenças discriminativas entre categorias.

**Máquinas de Vetores de Suporte (SVM)**

Introduzido por Cortes e Vapnik (1995), o SVM busca maximizar a margem entre hiperplanos separadores, sendo eficaz para problemas de alta dimensionalidade. No contexto BoF, o SVM é amplamente empregado devido: (i) à robustez contra *overfitting*; (ii) à eficiência para descritores histogramais; e (iii) à capacidade de lidar com dados ruidosos (BURGES, 1998).

**Doenças presentes no *Dataset Apple Tree Leaf***

O *dataset Apple Tree Leaf* contém quatro classes visualmente distintas (Zhong; Zhao, 2020). São elas:

*a) Alternaria leaf spot* (Mancha de Alternaria)

Causador: fungos do gênero *Alternaria* (principalmente *Alternaria alternata*).

Características visuais:

- Manchas circulares a irregulares, geralmente marrons escuras.
- Presença de anéis concêntricos, criando o aspecto de “alvo” ou “olho de boi”.
- Halo amarelado ao redor das lesões.
- Pode evoluir para necrose e perfurações.

Impacto:

- Reduz a fotossíntese.
- Pode causar queda prematura das folhas.


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218



A Figura 1 ilustra uma folha de macieira com sintomas característicos da doença em questão, evidenciando as manifestações visuais utilizadas como base para a análise e o reconhecimento no presente estudo.

**Figura 1**. Folha com a doença *alternaria leaf spot* (Mancha de Alternaria)



Fonte: https://www.kaggle.com/datasets/nirmalsankalana/apple-tree-leaf-disease-dataset. Acesso em 13. nov. 2025.

*b) Brown spot* (Mancha Marrom)

Causador: diversos fungos, variando por espécie vegetal (ex.: *Septoria, Helminthosporium, Phloeospora*).

Características visuais:

- Manchas marrons pequenas que aumentam gradualmente.
- Formato geralmente circular ou oval.
- Podem apresentar um centro claro e bordas mais escuras.
- Em estágios avançados, formam áreas necrosadas maiores.

Impacto:

- Redução da área fotossintética.
- Em plantas jovens, pode comprometer o crescimento.

A Figura 2 apresenta uma amostra representativa de *Brown spot*, destacando manchas de coloração marrom, tipicamente circulares ou ovais, que se expandem gradualmente e podem culminar em necrose, condição associada à redução da área fotossintética e, em estágios precoces de desenvolvimento, à limitação do crescimento da planta.


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

**Figura 2**. Folha com a doença *Brown spot* (Mancha Marrom)



Fonte: https://www.kaggle.com/datasets/nirmalsankalana/apple-tree-leaf-disease-dataset. Acesso em 13. nov. 2025.

*c) Gray spot* (Mancha Cinzenta)

Causador: frequentemente *Stemphylium spp., Cercospora spp.* ou *Pyricularia spp.,* variando conforme a cultura.

Características visuais:

- Manchas inicialmente acastanhadas, que evoluem para cinza claro.
- Bordas mais escuras ou avermelhadas.
- Textura seca e aspecto “queimado”.
- Em alguns casos, formação de pequenas perfurações.

Impacto:

- Afeta severamente folhas mais velhas.
- Pode favorecer infecções secundárias.

A Figura 3 ilustra uma amostra representativa de *Gray spot*, na qual se observam manchas que progridem de acastanhadas para cinza claro, com bordas mais escuras ou avermelhadas e textura seca, podendo ocorrer pequenas perfurações; esse padrão é consistente com a descrição de maior severidade em folhas mais velhas e potencial predisposição a infecções secundárias.


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

**Figura 3**. Folha com a doença *Gray spot* (Mancha Cinzenta)



Fonte: https://www.kaggle.com/datasets/nirmalsankalana/apple-tree-leaf-disease-dataset. Acesso em 13. nov. 2025.

*d) Rust* (Ferrugem)

Causador: fungos da ordem *Pucciniales* (ex.: *Puccinia spp*.).

Características visuais:

- Pústulas (pontinhos) alaranjadas, avermelhadas ou amarelas.
- Aspecto semelhante a “ferrugem metálica”, motivo do nome.
- As pústulas podem liberar esporos de coloração intensa.
- Em estágios avançados, causam manchas escuras e necrose.

Impacto:

- Altamente disseminável pelo vento.
- Uma das doenças foliares mais agressivas para culturas agrícolas.

A Figura 4 ilustra uma amostra representativa de *Rust* (ferrugem), destacando pústulas intensamente pigmentadas e potencial liberação de esporos, características associadas à elevada capacidade de disseminação pelo vento e à agressividade do patógeno em sistemas agrícolas.

**Figura 4**. Folha com a doença *Rust* (Ferrugem)



Fonte: https://www.kaggle.com/datasets/nirmalsankalana/apple-tree-leaf-disease-dataset. Acesso em 13. nov. 2025.


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

## MÉTODOS

### Base de dados

As imagens foram coletadas do *dataset Apple Tree Leaf*, organizado em pastas por classe. Assim, o conjunto de dados utilizado neste estudo é composto por imagens reais de folhas de macieira coletadas em quatro estações experimentais vinculadas à Northwest *University of Agriculture and Forestry Science and Technology*, localizada na região noroeste da China. As fotografias foram obtidas por meio de um dispositivo móvel *Glory* V10, o que garante variabilidade natural nas condições de captura.

As folhas foram registradas tanto em ambiente controlado de laboratório (aproximadamente 51,9% das imagens) quanto em áreas de cultivo reais (cerca de 48,1%), contemplando diferentes condições de iluminação, clima e horários do dia. Essa diversidade confere ao *dataset* características importantes para a avaliação de algoritmos de visão computacional, uma vez que inclui variações naturais encontradas em cenários agrícolas reais. O conjunto contempla quatro doenças de grande relevância agronômica: *Alternaria leaf spot*, *Gray spot*, *Brown spot* e *Rust*, apresentadas em distintos níveis de severidade, o que contribui para a heterogeneidade visual e para a robustez dos experimentos de classificação. O carregamento foi realizado em MATLAB via:

```
treinamentoset = imageSet('DATA-SET-APPLE-TREE-LEAF','recursive');
```

### Balanceamento e divisão

Para evitar viés de treinamento, utilizou-se:

```
mintreinamento = min([treinamentoset.Count]);
treinamento = partition(treinamentoset, mintreinamento, 'randomize');
[treino, teste] = partition(treinamento, 0.6, 'randomize');
```

Dessa forma, o conjunto de dados foi balanceado, mantendo-se o mesmo número de imagens em cada classe, e posteriormente particionado em 60% para treinamento e 40% para teste, assegurando uma avaliação comparável do desempenho do modelo entre as categorias.

### Variação do vocabulário visual

O experimento central consistiu em variar o vocabulário de 100 a 3000 em incrementos de 200:

```
bag = bagOfFeatures(treino, 'VocabularySize', k, 'PointSelection', 'Detector');
```

onde:

$k \in \{100, 300, 700, 1000, 3000\}$

Para cada vocabulário:


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

1. o BoF é recriado;
2. um novo SVM é treinado;
3. o modelo é avaliado;
4. a acurácia é armazenada.

**Classificação e avaliação**

O SVM multiclasses foi treinado via:

```
categoryClassifier = trainImageCategoryClassifier(treino, bag);
```

A avaliação usou matriz de confusão:

```
[confMat,~,~,~] = evaluate(categoryClassifier, teste);
acc = mean(diag(confMat));
```

A acurácia global foi estimada a partir da soma dos elementos da diagonal principal da matriz de confusão (isto é, o total de predições corretas), normalizada pelo número total de amostras, conforme procedimento adotado na literatura (Zhong; Zhao, 2020).

## RESULTADOS E DISCUSSÃO

Para a avaliação experimental, utilizou-se um computador equipado com 12 GB de memória RAM e processador Intel Core i7 (3ª geração), com implementação e execução dos procedimentos no ambiente MATLAB.

No primeiro experimento, adotou-se um vocabulário visual com 100 palavras. Nessa configuração, o modelo atingiu acurácia média global de 73% e tempo de processamento de 53,54 s. A Figura 5 apresenta a matriz de confusão correspondente, permitindo a análise do desempenho por classe e dos padrões de erro associados ao vocabulário reduzido.

**Figura 5**. Matriz de confusão para um vocabulário = 100

| | PREDICTED | | | | |
|---|---|---|---|---|---|
| KNOWN | Alternaria leaf spot | Brown spot | Gray spot | Healthy leaf | Rust |
| Alternaria leaf spot | 0.88 | 0.02 | 0.00 | 0.02 | 0.07 |
| Brown spot | 0.02 | 0.85 | 0.00 | 0.05 | 0.08 |
| Gray spot | 0.07 | 0.23 | 0.06 | 0.38 | 0.26 |
| Healthy leaf | 0.01 | 0.03 | 0.01 | 0.94 | 0.00 |
| Rust | 0.02 | 0.05 | 0.00 | 0.01 | 0.92 |

Fonte: elaboração própria, a partir dos resultados experimentais (MATLAB R2023b).

A análise da matriz de confusão (Figura 5) indica desempenho elevado para as classes *Healthy leaf* (94%), *Rust* (92%), *Alternaria leaf spot* (88%) e *Brown spot* (85%), evidenciando adequada capacidade discriminativa do modelo para essas categorias. Em contrapartida, observa-se degradação acentuada na classe Gray spot, cuja taxa de reconhecimento foi de apenas 6%, indicando dificuldade substantiva de separação dessa classe no espaço de atributos gerado pelo


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218



vocabulário reduzido. Esse comportamento sugere maior sobreposição das assinaturas visuais da *Gray spot* com outras categorias, implicando limitações de generalização do modelo sob baixa granularidade de quantização.

No segundo experimento, adotou-se um vocabulário visual com 300 palavras. Nessa configuração, o modelo alcançou acurácia média global de 77% e tempo de processamento de 50,4 s. A Figura 6 apresenta a matriz de confusão correspondente, possibilitando a avaliação do desempenho por classe e a identificação dos principais padrões de confusão.

**Figura 6.** Matriz de confusão para um vocabulário = 300

| | PREDICTED | | | | |
|---|---|---|---|---|---|
| KNOWN | Alternaria leaf spot | Brown spot | Gray spot | Healthy leaf | Rust |
| Alternaria leaf spot | 0.88 | 0.01 | 0.03 | 0.03 | 0.03 |
| Brown spot | 0.01 | 0.79 | 0.08 | 0.03 | 0.08 |
| Gray spot | 0.03 | 0.12 | 0.31 | 0.22 | 0.31 |
| Healthy leaf | 0.03 | 0.01 | 0.01 | 0.94 | 0.00 |
| Rust | 0.01 | 0.01 | 0.00 | 0.03 | 0.94 |

Fonte: elaboração própria, a partir dos resultados experimentais (MATLAB R2023b).

A partir da matriz de confusão apresentada na Figura 6, observa-se desempenho consistente para as classes *Alternaria leaf spot* (88%), *Healthy leaf* (94%) e *Rust* (94%), enquanto *Brown spot* apresentou taxa de reconhecimento moderada (79%). Em contraste, a classe Gray spot permaneceu como a principal fonte de erro do modelo, com 31% de acerto, indicando persistente sobreposição de padrões visuais com outras categorias mesmo com o aumento do vocabulário para 300 palavras.

No terceiro experimento, adotou-se um vocabulário visual com 700 palavras. Nessa configuração, o modelo apresentou acurácia média global de 77% e tempo de processamento de 51,31 s. A Figura 7 apresenta a matriz de confusão correspondente, permitindo a análise do desempenho por classe e dos principais padrões de confusão.

**Figura 7**. Matriz de confusão para um vocabulário = 700

| | PREDICTED | | | | |
|---|---|---|---|---|---|
| KNOWN | Alternaria leaf spot | Brown spot | Gray spot | Healthy leaf | Rust |
| Alternaria leaf spot | 0.93 | 0.00 | 0.01 | 0.01 | 0.05 |
| Brown spot | 0.02 | 0.84 | 0.07 | 0.02 | 0.05 |
| Gray spot | 0.01 | 0.09 | 0.29 | 0.34 | 0.27 |
| Healthy leaf | 0.00 | 0.02 | 0.08 | 0.90 | 0.00 |
| Rust | 0.01 | 0.00 | 0.03 | 0.05 | 0.91 |

Fonte: elaboração própria, a partir dos resultados experimentais (MATLAB R2023b).

A análise da matriz de confusão apresentada na Figura 7 evidencia desempenho elevado *para Alternaria leaf spot* (93%), *Brown spot* (84%), *Healthy leaf* (90%) e *Rust* (91%), indicando boa capacidade discriminativa do modelo para essas classes. Em contraste, a classe Gray spot


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

manteve desempenho substancialmente inferior (29%), reiterando a dificuldade do sistema em separar essa categoria das demais, possivelmente em razão de maior similaridade textural e sobreposição de padrões visuais no espaço de características gerado pelo BoF.

Dando sequência aos experimentos, na quarta simulação foi adotado um vocabulário visual com 1000 palavras. Nessa configuração, o modelo alcançou acurácia média global de 78% e tempo de processamento de 53,4 s. A Figura 8 apresenta a matriz de confusão correspondente, permitindo avaliar o desempenho por classe e os principais padrões de confusão.

**Figura 8.** Matriz de confusão para um vocabulário = 1000

| | | PREDICTED | | | |
|---|---|---|---|---|---|
| KNOWN | Alternaria leaf spot | Brown spot | Gray spot | Healthy leaf | Rust |
| Alternaria leaf spot | 0.92 | 0.00 | 0.00 | 0.05 | 0.03 |
| Brown spot | 0.01 | 0.84 | 0.07 | 0.05 | 0.03 |
| Gray spot | 0.00 | 0.05 | 0.35 | 0.37 | 0.23 |
| Healthy leaf | 0.02 | 0.02 | 0.01 | 0.93 | 0.01 |
| Rust | 0.00 | 0.01 | 0.13 | 0.01 | 0.85 |

Fonte: elaboração própria, a partir dos resultados experimentais (MATLAB R2023b).

A análise da Figura 8 indica desempenho elevado para *Alternaria leaf spot* (92%) e *Healthy leaf* (93%), bem como resultados consistentes para *Brown spot* (84%) e *Rust* (85%), sugerindo adequada separabilidade dessas classes no espaço de características quantizado pelo vocabulário de 1000 palavras. Em contraste, a classe *Gray spot* apresentou desempenho substancialmente inferior (35%), evidenciando persistente dificuldade do modelo em discriminar essa categoria em relação às demais, o que aponta para maior sobreposição de padrões visuais e maior propensão a confusões interclasse mesmo com o aumento do vocabulário.

No quinto experimento, adotou-se um vocabulário visual com 3000 palavras. Nessa configuração, o modelo apresentou acurácia média global de 77% e tempo de processamento de 80,7 s. A Figura 9 apresenta a matriz de confusão correspondente, possibilitando a análise do desempenho por classe e dos padrões de confusão associados ao aumento do vocabulário.

**Figura 9.** Matriz de confusão para um vocabulário = 3000

| | | PREDICTED | | | |
|---|---|---|---|---|---|
| KNOWN | Alternaria leaf spot | Brown spot | Gray spot | Healthy leaf | Rust |
| Alternaria leaf spot | 0.91 | 0.00 | 0.01 | 0.05 | 0.03 |
| Brown spot | 0.02 | 0.74 | 0.10 | 0.02 | 0.10 |
| Gray spot | 0.03 | 0.07 | 0.42 | 0.31 | 0.16 |
| Healthy leaf | 0.01 | 0.05 | 0.05 | 0.88 | 0.01 |
| Rust | 0.01 | 0.00 | 0.06 | 0.01 | 0.92 |

Fonte: elaboração própria, a partir dos resultados experimentais (MATLAB R2023b).

A matriz de confusão apresentada na Figura 9 sintetiza o desempenho do classificador BoF+SVM nas cinco classes avaliadas. Observa-se desempenho elevado para *Rust* (92%) e


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

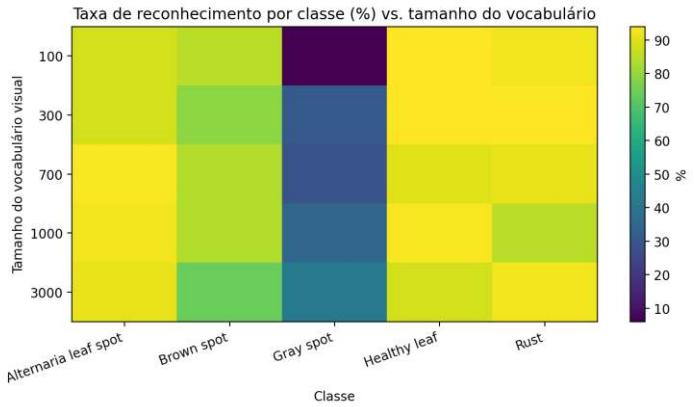
*Alternaria leaf spot* (91%), bem como resultado consistente para *Healthy leaf* (88%), indicando adequada separabilidade dessas categorias no espaço de características quantizado pelo vocabulário de 3000 palavras. A classe *Brown spot* apresentou desempenho intermediário, com 74% de acerto, sugerindo maior sobreposição parcial de padrões visuais em relação às classes de maior desempenho.

Em contrapartida, a classe Gray spot permaneceu como a principal fonte de erro do modelo, com taxa de reconhecimento de 42%. Adicionalmente, verificam-se confusões expressivas dessa classe com *Healthy leaf* (31%) e, em menor magnitude, com *Rust* (16%). Esse padrão indica que atributos visuais característicos da *Gray spot* — particularmente componentes de textura e variações de coloração — podem apresentar similaridade com outras categorias, limitando a separação no espaço de atributos induzido pelo BoF e contribuindo para erros de classificação mesmo sob vocabulário mais granular.

**Análise comparativa do efeito do tamanho do vocabulário visual**

As Figuras 10, 11 e 12 sintetizam, de forma comparativa, o impacto do tamanho do vocabulário visual sobre o desempenho e o custo computacional do pipeline BoF+SVM. Observa-se na Figura 10 que a acurácia global evolui de 73% (vocabulário = 100) para um patamar próximo de saturação entre 700 e 1000 palavras (77–78%), sugerindo que vocabulários intermediários oferecem o melhor equilíbrio entre poder discriminativo e capacidade de generalização. Em contrapartida, a Figura 11 evidencia que o aumento do vocabulário não implica ganhos proporcionais de desempenho, mas eleva o custo de processamento, com crescimento expressivo do tempo no vocabulário 3000 (80,7 s), caracterizando um cenário de retorno marginal em termos de acurácia frente ao custo computacional.

**Figura 10.** Acurácia geral (%) em função do tamanho do vocabulário visual (escala log)




Fonte: elaboração própria, a partir dos resultados experimentais (MATLAB R2023b).

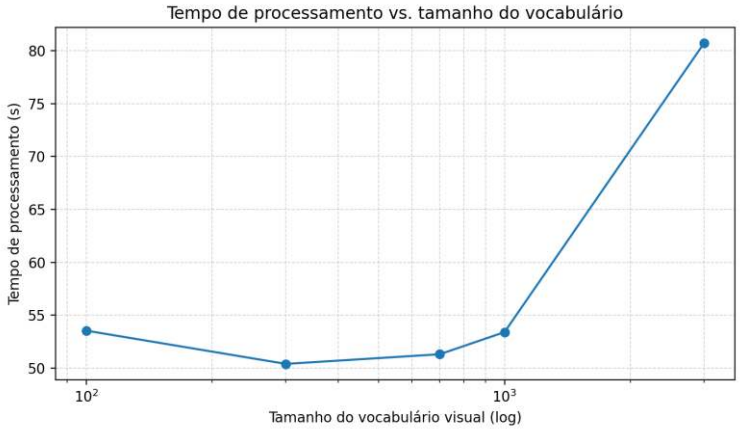

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218



A Figura 10 evidencia que o aumento do tamanho do vocabulário visual produz ganhos de acurácia global apenas até um certo limiar: há uma elevação inicial do desempenho ao sair de 100 palavras (73%) para valores intermediários, seguida por um platô entre 700 e 1000 palavras (≈ 77–78%), sem melhora adicional relevante ao ampliar o vocabulário para 3000 (≈ 77%). Esse comportamento indica retorno marginal decrescente do aumento de granularidade da quantização, sugerindo que vocabulários intermediários oferecem o melhor compromisso entre expressividade da representação e capacidade de generalização do modelo BoF+SVM no conjunto avaliado.

**Figura 11**. Tempo de processamento (s) em função do tamanho do vocabulário visual (escala log)

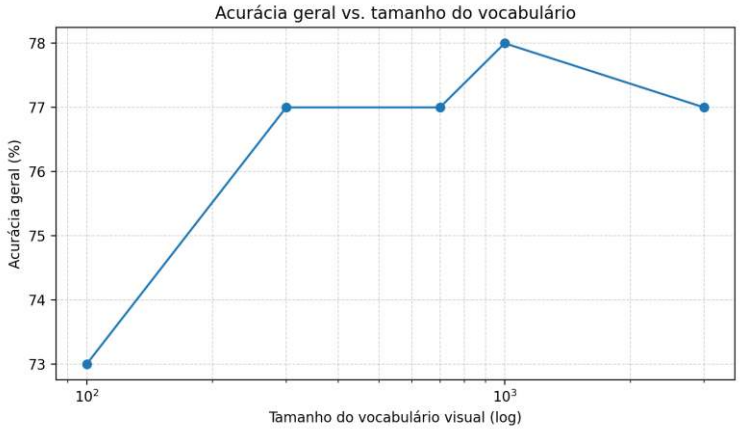



Fonte: elaboração própria, a partir dos resultados experimentais (MATLAB R2023b).

A Figura 11 demonstra que o tempo de processamento cresce com o aumento do tamanho do vocabulário visual, refletindo o maior custo computacional associado à quantização e ao manuseio de histogramas mais complexos. Observa-se que, para vocabulários entre 100 e 1000 palavras, o tempo permanece relativamente estável (≈ 50–54 s), enquanto a configuração com 3000 palavras impõe um aumento expressivo, alcançando 80,7 s. Esse resultado evidencia que vocabulários muito extensos elevam significativamente o custo de execução, sem garantir ganhos proporcionais de desempenho, reforçando a necessidade de selecionar tamanhos intermediários quando se busca melhor relação entre acurácia e eficiência computacional.

ISSN: 2675-6218 - RECIMA21
Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

**Figura 12.** Mapa de calor da taxa de reconhecimento por classe (%) em função do tamanho do vocabulário visual




Fonte: elaboração própria, a partir dos resultados experimentais (MATLAB R2023b).

A análise por classe, apresentada na Figura 12, reforça a interpretação anterior ao evidenciar que o desempenho médio é fortemente condicionado por assimetrias entre categorias. Enquanto *Alternaria leaf spot, Healthy leaf* e *Rust* mantêm taxas elevadas e relativamente estáveis ao longo das configurações (≈ 85–94%), a classe *Gray spot* permanece como o principal fator limitante do modelo, ainda que apresente melhora progressiva com o aumento do vocabulário (6% → 31% → 29% → 35% → 42%). Esse padrão indica que a maior granularidade da quantização contribui para capturar variações úteis dessa classe, porém não elimina o problema de separabilidade, o que é compatível com a maior sobreposição de atributos visuais (especialmente textura e coloração) entre *Gray spot* e classes como *Healthy leaf* e *Rust*, já observada nas matrizes de confusão. Em conjunto, os resultados resumidos nas Figuras 10–12 corroboram a conclusão de que vocabulários intermediários (especialmente 700–1000) maximizam a relação desempenho–custo do pipeline proposto, ao passo que vocabulários muito extensos tendem a aumentar a complexidade computacional sem ganhos equivalentes de acurácia global, mantendo Gray spot como a classe mais desafiadora para a abordagem BoF+SVM.

## CONSIDERAÇÕES

Os resultados deste estudo demonstram que o pipeline clássico *Bag of Features* (BoF) + *Support Vector Machine* (SVM) constitui uma alternativa tecnicamente consistente para o reconhecimento de doenças foliares em macieira, particularmente em cenários com bases de tamanho moderado e restrições computacionais. A avaliação sistemática do tamanho do vocabulário visual (100, 300, 700, 1000 e 3000 palavras) evidenciou um comportamento típico de ganhos marginais decrescentes: observou-se incremento de desempenho ao sair de vocabulários reduzidos, seguido por um patamar de saturação em configurações intermediárias. Em termos práticos, os vocabulários entre 700 e 1000 apresentaram o melhor compromisso entre poder


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

discriminativo e capacidade de generalização, ao passo que o aumento para 3000 não resultou em ganho proporcional de acurácia global e impôs custo computacional significativamente maior, caracterizando um trade-off desfavorável para aplicações operacionais.

A análise por classe reforça que o desempenho agregado do classificador é condicionado por assimetrias relevantes entre categorias. Enquanto Alternaria *leaf spot, Brown spot, Rust* e *Healthy leaf* apresentaram taxas de reconhecimento elevadas e relativamente estáveis ao longo das configurações, a classe Gray spot permaneceu como o principal fator limitante do modelo, com persistência de confusões com classes visualmente próximas. Esse resultado sugere que, no espaço de atributos induzido pelo BoF, os padrões característicos da Gray spot apresentam maior sobreposição com outras categorias, restringindo a separabilidade mesmo sob vocabulários mais granulares. Assim, embora o aumento do vocabulário contribua para capturar variações mais finas e elevar parcialmente o desempenho dessa classe, a evidência empírica indica que a limitação não é apenas de granularidade, mas também de representatividade do descritor e da própria ambiguidade visual entre estados patológicos e não patológicos em condições reais de aquisição.

Do ponto de vista metodológico e aplicado, os achados corroboram a relevância de abordagens clássicas de visão computacional como soluções interpretáveis, reprodutíveis e computacionalmente viáveis para diagnóstico foliar automatizado, especialmente quando se busca desempenho robusto sem depender de grandes volumes de dados rotulados e infraestrutura de alto custo. Como direções de pesquisa, recomenda-se aprofundar estratégias especificamente orientadas à classe Gray spot, incluindo: (i) pré-processamento e segmentação para reduzir interferências do fundo e realçar regiões lesionadas; (ii) incorporação de descritores complementares (p.ex., texturais e cromáticos) e esquemas de codificação mais informativos; (iii) exploração de modelos híbridos, como CNNs leves ou mecanismos de extração profunda acoplados a classificadores lineares, preservando viabilidade computacional; e (iv) ampliação e refinamento do conjunto de dados, com ênfase em variabilidade intraclasse e padronização de anotações para reduzir ruído de rótulo. Em síntese, conclui-se que a abordagem BoF+SVM, quando calibrada por vocabulários intermediários, oferece uma base sólida para sistemas de apoio ao diagnóstico em agricultura de precisão, ao mesmo tempo em que explicita limites claros de separabilidade para classes com assinaturas visuais menos pronunciadas, delineando um caminho objetivo para aprimoramentos futuros.

## REFERÊNCIAS

ABADE, André; FERREIRA, Paulo Afonso; VIDAL, Flavio de Barros. Plant diseases recognition on images using convolutional neural networks: a systematic review. **Computers and Electronics in Agriculture**, v. 185, p. 106125, jun. 2021. DOI: 10.1016/j.compag.2021.106125.


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

BARBEDO, J. G. A. A review on the main challenges in automatic plant disease identification based on visible range images. **Biosystems Engineering**, v. 144, p. 52-60, 2016. DOI: 10.1016/j.biosystemseng.2016.01.017

BURGES, Christopher J. C. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge **Discovery**, v. 2, p. 121–167, 1998. DOI: 10.1023/A:1009715923555. Disponível em: https://doi.org/10.1023/A:1009715923555. Acesso em: 3 fev. 2026.

CORTES, Corinna; VAPNIK, Vladimir. Support-vector networks. **Machine Learning**, v. 20, n. 3, p. 273–297, 1995. DOI: 10.1007/BF00994018. Disponível em: https://doi.org/10.1007/BF00994018. Acesso em: 3 fev. 2026

CSURKA, Gabriella; DANCE, Christopher R.; FAN, Lixin; WILLAMOWSKI, Jutta; BRAY, Cédric. Visual categorization with bags of keypoints. *In:* **EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV)**, 8., 2004, Prague. Proceedings… Workshop on Statistical Learning in Computer Vision. Xerox Research Centre Europe (XRCE), 2004. p. 1–22. Disponível em: https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/csurka-eccv-04.pdf. Acesso em: 1 fev. 2026.

FERENTINOS, K. P. Deep learning models for plant disease detection and diagnosis. **Computers and Electronics in Agriculture**, v. 145, p. 311–318, 2018. DOI: 10.1016/j.compag.2018.01.009

HUANG, Yongzhen; WU, Zifeng; WANG, Liang; TAN, Tieniu. Feature coding in image classification: a comprehensive study. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 36, n. 3, p. 493–506, 2014. DOI: 10.1109/TPAMI.2013.113. Disponível em: https://doi.org/10.1109/TPAMI.2013.113. Acesso em: 2 fev. 2026.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015. DOI: 10.1038/nature14539.

MOHANTY, S. P.; HUGHES, D. P.; SALATHÉ, M. Using deep learning for image-based plant disease detection. **Frontiers in Plant Science**, v. 7, art. 1419, 2016. DOI: 10.3389/fpls.2016.01419

NOWAK, Eric; JURIE, Frédéric; TRIGGS, Bill. Sampling strategies for bag-of-features image classification. *In:* **EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV)**, 9., 2006. Proceedings… Berlin: Springer, 2006. p. 490–503. (Lecture Notes in Computer Science). DOI: 10.1007/11744085_38. Disponível em: https://doi.org/10.1007/11744085_38. Acesso em: 3 fev. 2026.

PAWARA, Pornntiwa; OKAFOR, Emmanuel; SURINTA, Olarik; SCHOMAKER, Lambertus; WIERING, Marco. Comparing Local Descriptors and Bags of Visual Words to Deep Convolutional Neural Networks for Plant Recognition. *In:* **INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION APPLICATIONS AND METHODS (ICPRAM)**, 6., 2017, Porto. Proceedings… [S. l.]: SciTePress, 2017. p. 479–486. DOI: 10.5220/0006196204790486. Disponível em: https://www.scitepress.org/papers/2017/61962/61962.pdf. Acesso em: 1 fev. 2026.

PIRES, Rillian Diello Lucas; GONÇALVES, Diogo Nunes; ORUÊ, Jonatan Patrick Margarido; KANASHIRO, Wesley Eiji Sanches; RODRIGUES JR., Jose F.; MACHADO, Bruno Brandoli; GONÇALVES, Wesley Nunes. Local descriptors for soybean disease recognition. **Computers and Electronics in Agriculture**, v. 125, p. 48–55, 2016. DOI: 10.1016/j.compag.2016.04.032.

YANG, Jun; JIANG, Yu-Gang; HAUPTMANN, Alexander G.; NGO, Chong-Wah. Evaluating bag-of-visual-words representations in scene classification. *In:* **INTERNATIONAL WORKSHOP ON**


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218



**MULTIMEDIA INFORMATION RETRIEVAL (MIR)**, 2007. Proceedings… New York: ACM, 2007. p. 197–206. DOI: 10.1145/1290082.1290111. Disponível em: https://doi.org/10.1145/1290082.1290111. Acesso em: 3 fev. 2026.

ZHONG, Yong; ZHAO, Ming. Research on deep learning in apple leaf disease recognition. **Computers and Electronics in Agriculture**, v. 168, art. 105146, 2020. DOI: 10.1016/j.compag.2019.105146. Disponível em: https://doi.org/10.1016/j.compag.2019.105146. Acesso em: 3 fev. 2026.


Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.