

COMPARAÇÃO DE ESTRATÉGIAS DE TREINAMENTO PARA OTIMIZAÇÃO DE MODELOS DE CLASSIFICAÇÃO DE AFÍDEOS

Autor Principal: Brenda Slongo Taca
197402@upf.br

Coautores: Brenda Slongo Taca; Henrique Alberto Bertolin Costa; Douglas Lau

Orientador: Rafael Rieder

Subevento: Ciências Exatas e da Terra

Área: Mostra de Iniciação Científica

Tipo de Trabalho: Resumo

Introdução: Conforme Lins et al. (2020), o AphidCV é um software para detecção, contagem, classificação e mensuração de afídeos, popularmente conhecidos como pulgões. Rodriguez e Rieder (2020) propuseram uma arquitetura de rede neural própria de classificação do AphidCV, que resultou em modelos com significativa melhora no desempenho computacional e na predição de afídeos. Com a integração do software como serviço na plataforma TrapSystem (Kirinus et al., 2023), e com o crescimento do dataset com novas imagens e a inclusão de novas espécies, é essencial criar novos modelos e atualizar os já existentes, considerando novas estratégias de treinamento para melhorar métricas de avaliação e, conseqüentemente, a predição correta. Com isso em mente, o objetivo deste trabalho é apresentar uma comparação de estratégias de treinamento para otimização de modelos de classificação de afídeos da arquitetura do AphidCV.

Metodologia: Primeiramente, organizou-se o dataset disponibilizado e rotulado pela Embrapa Trigo, parceria do projeto AphidCV, armazenado em pasta compartilhada na nuvem. O banco vigente dispõe um total de 148.340 imagens recortadas de insetos de seis espécies: *Rhopalosiphum padi* (Rp), *Schizaphis graminum* (Sg), *Metopolophium dirhodum* (Md), *Sitobion avenae* (Sa), *Myzus persicae* (Mp) e *Brevicoryne brassicae* (Bb). Cada espécie é categorizada em quatro classes: alados (afídeos com asas), ápteros (afídeos adultos), ninfas (afídeos jovens) e falsos (exúvias e detritos). Nessa etapa, observou-se que o dataset estava desbalanceado em relação ao total de imagens por classe e por espécie. De acordo com Jonhson e Khoshgoftaar (2019), apesar do desbalanceamento de classe ser inerente a muitas aplicações do mundo real, conjuntos de dados altamente desequilibrados podem ocasionar viés em relação à classe majoritária e ignorar a classe minoritária. Logo, para este estudo, considerou-se a quantidade da classe minoritária: 2.000 imagens. Criou-se um clone local do banco e, aleatoriamente, foram selecionadas 2.000 imagens por classe para cada espécie, totalizando 48.000 imagens. O passo seguinte foi definir estratégias de treinamento para otimização e criação de novos modelos para cada espécie de afídeo. Foram criadas quatro versões da arquitetura de rede neural, variando apenas hiperparâmetros de batch_

size e patience. Esses parâmetros representam, respectivamente, a quantidade de amostras de dados que são transferidas para a rede e o controle de parada antecipada de treinamento (caso a acurácia não melhore em determinado intervalo de épocas). O Teste 1 considerou batch_size de 100 amostras, e patience de 5 épocas; o Teste 2 considerou 100 e 10; o Teste 3 considerou 32 e 5; e o Teste 4 considerou 32 e 10. Em virtude do número restrito de imagens por classe, adotou-se também uma estratégia de preparação de dados. Aplicou-se filtro de normalização no pré-processamento de todas as imagens de entrada. No treinamento, adotou-se técnicas de aumento de dados para maior variabilidade de entrada, considerando transformações geométricas (rotação, espelhamento e redimensionamento) e filtros convolucionais (borramento, nitidez, relevo, dilatação, erosão, equalização de histograma e variação de brilho e contraste), com probabilidade de aplicação de 25% por época. A Figura 1 apresenta os resultados. Observa-se que a melhor estratégia, na média, é o Teste 4, com dois dos melhores modelos (Bb e Rp). Porém, três dos melhores modelos gerados estão na estratégia do Teste 2 (Md, Mp e Sa). Pode-se dizer que o melhor parâmetro patience é 10 épocas e, para um conjunto balanceado com quantidade reduzida de imagens por classe, o tamanho do batch_size não é significativo. No caso do modelo Sg, que apresentou acurácia inferior a 80%, notou-se no dataset que haviam algumas poucas imagens com características diferentes da maioria - o que pode ter influenciado o resultado.

Conclusão: A comparação mostrou que a estratégia de batch_size de 32 e patience de 10 obteve o melhor desempenho. Porém, para evitar o overfitting e para generalizar os modelos, é fundamental treinar com um conjunto maior de imagens por espécie. Os próximos passos estão relacionados ao aumento do dataset e à comparação com arquiteturas consolidadas do estado da arte.

Referências: JOHNSON, J. M.; KHOSHGOFTAAR, T. M. Survey on deep learning with class imbalance. Journal of Big Data, v. 6, n. 1, p. 1-54, 2019.
KIRINUS, N. W. et al. AphidCV 3.0: Abordagem Web Integrada a uma Plataforma de Monitoramento de Insetos. In: III Congresso de Tecnologia da Informação do IFSul Passo Fundo. IFSul, 2023. In press.
LINS, E. A. et al. A method for counting and classifying aphids using computer vision. Computers and Electronics in Agriculture, v. 169, p. 105200, 2020.
RODRIGUEZ, J. P. M.; RIEDER, R. AphidCV 2.0: uma nova abordagem de classificação, contagem e mensuração de afídeos. In: Anais Estendidos do XXXIII Conference on Graphics, Patterns and Images. SBC, 2020. p. 159-162.

Espécie	Teste 1 Batch size=100; Patience=5		Teste 2 Batch size=100; Patience=10		Teste 3 Batch size=32; Patience=5		Teste 4 Batch size=32; Patience=10	
	Early stopping	Resultado	Early stopping	Resultado	Early stopping	Resultado	Early stopping	Resultado
<i>Brevicoryne brassicae</i>	46	77,96%	78	80,64%	40	79,93%	150	84,48%
<i>Metopolophium dirhodum</i>	42	83,50%	138	87,19%	62	83,69%	123	86,63%
<i>Myzus persicae</i>	22	65,57%	139	85,36%	104	80,59%	150	84,93%
<i>Rhopalosiphum padi</i>	66	87,17%	59	80,81%	51	79,94%	62	87,63%
<i>Schizaphis graminum</i>	64	78,62%	123	78,06%	42	69,00%	133	72,88%
<i>Sitobion avenae</i>	38	80,31%	150	90,00%	69	82,13%	150	88,31%
Média	46	78,86%	115	83,68%	61	79,21%	128	84,14%