

## APRENDIZADO PROFUNDO EMBARCADO EM ARMADILHAS INTELIGENTES PARA A IDENTIFICAÇÃO DE INSETOS-PRAGA

LUCAS COUTINHO FREITAS<sup>1</sup>; VALTER ALBERTO MELGAREJO  
MARTINS<sup>1</sup>; PAULO ROBERTO FERREIRA JÚNIOR<sup>1</sup>; LISANE BRISOLARA DE  
BRISOLARA<sup>1</sup>

<sup>1</sup>Universidade Federal de Pelotas – Grupo de Pesquisa em Engenharia de Sistemas Ciber-Físicos  
{lcfreitas,vammartins,paulo.ferreira,lisane}@inf.ufpel.edu.br

### 1. INTRODUÇÃO

Redes de sensores sem fio (RSSF) compostas por armadilhas eletrônicas podem ser empregadas para o monitoramento de pragas no campo. A armadilha atrai insetos-praga e através do processamento das imagens capturadas pelo seu sensor, estes serão identificados e contabilizados (FREITAS et al., 2019; FREITAS et al., 2018). A partir da densidade de pragas identificadas por um ou mais sensores, pode-se prever a eclosão de infestações, permitindo a utilização de técnicas de controle de pragas mais racionais.

Dada as limitações energéticas dos nodos sensores, são empregados protocolos de roteamento eficientes e técnicas de distribuição de carga. Além disso, a aplicação deve ser desenvolvida considerando o custo de processamento e transmissão de dados de forma a maximizar o tempo de vida da rede.

Nossa aplicação emprega imagens para identificação dos insetos - cada imagem tem em torno de 8MB. O envio destas imagens pela rede tem um custo de comunicação proibitivo, visto que os módulos de rádio frequência (RF) que visam menor consumo e longo alcance geralmente são limitados a baixas taxas de transmissão - em torno de 50Kb a 250 Kb por segundo, como discutido em (FREITAS et al., 2019). Logo, assumindo uma rede que contenha dez nodos e um deles sendo o nó concentrador, e que a cada salto que os pacotes fazem é reduzida a taxa de transmissão, fica evidente que a transmissão dessas imagens se torna inviável. Esta limitação dos rádios motivou o processamento local das imagens das armadilhas. Consequentemente, será transmitido pelos nodos somente a contagem de insetos, *Ceratitis* e *Grapholita*, realizada na etapa de identificação. Assim, a aplicação envia um único pacote com 8 bytes por payload (2 inteiros) mais cabeçalhos inerentes ao protocolo.

No entanto, o processamento de imagens para a identificação dos insetos, também é uma tarefa custosa sendo necessário definir uma solução que apresente um baixo tempo de processamento e consequentemente baixo consumo energético. A identificação dos insetos praga pode ser dividida em duas partes: segmentação de regiões de interesse e a classificação dessas; ao final, é somado a quantidade de insetos presentes e são transmitidos esses valores para o nó concentrador.

Este trabalho aborda o desenvolvimento da etapa de classificação, utilizando redes neurais convolucionais (CNN, de Convolutional Neural Networks) para a classificação desses insetos. A solução foi embarcada juntamente com os modelos treinados em uma Raspberry Pi 3 modelo B junto da câmera Pi v2. A motivação para o uso desses modelos de Aprendizado Profundo é dada pela excelente performance desses em classificação de imagens, demonstrada em diversos trabalhos recentes (HE et al., 2016; XIE et al., 2017), inclusive para a detecção de insetos em armadilhas (WING; TAYLOR, 2016). Neste trabalho são

avaliadas diferentes arquiteturas de CNN para discutir o problema de acurácia versus latência, além de avaliar o impacto da compressão dos modelos treinados para acelerar o processo de inferência.

## 2. METODOLOGIA

Para a obtenção dos dados utilizados no treinamento dos classificadores, o segmentador proposto por REMBOSKI et al. (2018) foi empregado. A base de dados foi gerada segmentando todas as regiões de interesse de imagens feitas por armadilhas em campo. Cada imagem foi anotada por duas pessoas treinadas.

Diversos classificadores foram treinados através da biblioteca para desenvolvimento de modelos de aprendizado profundo Pytorch e várias arquiteturas de classificadores também foram avaliadas. Ao final, dois modelos foram quantizados com o intuito de reduzir o tempo e o custo de processamento de inferência. A principal métrica utilizada para a avaliação dos modelos foi a acurácia: taxa de acertos do modelo.

## 3. RESULTADOS E DISCUSSÃO

Foram selecionadas sete arquiteturas amplamente utilizadas na literatura para a classificação de imagens e que apresentam boa eficiência nos principais benchmarks de avaliação, tais como, ImageNet e COCO *dataset*. São elas: ResNext, ResNet18, DenseNet121, MNASNet, MobileNet, Inception e SqueezeNet. Primeiramente, foram descartadas algumas dessas arquiteturas que não eram viáveis, tendo em vista nossos recursos computacionais. Depois, este conjunto foi reduzido selecionando as arquiteturas que tiveram no máximo uma hora de treinamento. Assim, foram descartadas também as ResNext e Inception, que obtiveram um tempo de treinamento muito alto. A Tabela 1 apresenta o tempo de treinamento e acurácia obtidos pelas arquiteturas avaliadas com nenhuma transformação nos dados.

Arquitetura	Tempo [min]	Acurácia [%]
ResNext	720	indefinido
ResNet18	52	84,28
DenseNet121	53	84,75
MNASNet	23	80,11
MobileNet	22	84,91
Inception	241	83,01
SqueezeNet	23	67,65

Tabela 1. Tempo de treinamento e acurácia dos modelos.

Depois de selecionados, os modelos foram treinados utilizando técnicas para evitar sobre-ajuste, *augmentation*, e transferência de aprendizado para o modelo convergir mais rápido. Ao todo, foram treinados 180 modelos, 36 de cada

arquitetura. Desses, foram selecionados os que tiveram maior acurácia no conjunto de teste. Os resultados em acurácia de classificação no conjunto de teste são apresentados na Tabela 2, onde observa-se que o modelo ResNet18 obteve a maior acurácia no conjunto de teste.

	ResNet18	DenseNet121	SqueezeNet	MNASNet	MobileNet
Acurácia	90,72%	89,45%	86,69%	88,43%	87,06%

Tabela 2. Acurácia dos modelos no conjunto de teste.

Na sequência, foi avaliada a performance de todos os modelos quando embarcados no sistema computacional presente na armadilha. Visando a otimização da latência de inferência dos modelos, o modelo ResNet18 e MobileNet foram comprimidos através do método de quantização, aproximando os valores *float* para *int*, explorando os recursos de seções dos processadores dedicadas para operação com inteiros. O método tem como objetivo uma redução de tamanho do modelo de quatro vezes, e que acelere o tempo de inferência. A Tabela 3 resume os principais resultados para esses experimentos; modelos com a letra Q no início indicam modelos quantizados. A acurácia destes modelos comprimidos são de 88,5% e 78% para QResNet e QMobileNet, respectivamente. A aceleração obtida foi de  $2 \times$  e  $16 \times$ , respectivamente.

Modelo	Tamanho [MB]	Latência [s]
DenseNet121	58,56	$5,11 \pm 0,43$
MNASNet	16,85	$7,87 \pm 0,30$
MobileNet	13,50	$7,42 \pm 0,18$
ResNet18	66,67	$1,95 \pm 0,19$
SqueezeNet	3,81	$1,25 \pm 0,09$
QResNet	16,7	$0,97 \pm 0,27$
QMobileNet	3,47	$0,44 \pm 0,20$

Tabela 3. Espaço em memória e latência dos modelos.

Considerando os resultados obtidos, o modelo QResNet foi escolhido para nossa aplicação, visto que este modelo obteve cerca de 2% de degradação na acurácia e alcançou aceleração de duas vezes e uma compressão de 75% do tamanho do modelo.

#### 4. CONCLUSÕES

Neste trabalho foram avaliados diferentes modelos de CNN para a classificação de insetos-praga em pomares de fruticultura. Visando a redução do tempo de processamento local, dois dos nossos modelos foram comprimidos e

suas performances de classificação e latência foram avaliadas quando embarcados em nosso sistema computacional. Com base nos experimentos realizados, o modelo QResNet foi eleito para a nossa aplicação devido a sua acurácia e a compressão alcançada.

O trabalho desenvolvido trouxe resultados e contribuições científicas, sendo publicado no principal simpósio brasileiro da área (MARTINS et al., 2019). Além disso, o trabalho desenvolvido tem potencial para gerar produtos ou patentes e envolve parceria com uma empresa na área de agronegócios e agricultura de precisão.

Para trabalhos futuros, planeja-se medir o impacto energético real de cada otimização empregada. Além disso, será estimado o consumo energético individual das armadilhas, e através dos dados empíricos, caracterizar simuladores, permitindo o estudo de casos mais complexos onde a rede é composta por centenas de nodos.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

DING, W. and TAYLOR, G.. Automatic moth detection from trap images for pest management. **Computers and Electronics in Agriculture**, vol.123, pp. 17–28, 04 2016.

FREITAS, L. C.; MARTINS, V. A. M.; FERREIRA, P. R.; BRISOLARA, L. B. B.. Modelagem de Aplicações de Sensoriamento Visando Mapeamento Adaptativo em Redes de Sensores sem Fio. In: Congresso de Inovação Tecnológica da Universidade Federal de Pelotas, II., 2018. **Anais...** Pelotas, 2018.

FREITAS, L. C.; MARTINS, V. A. M.; FERREIRA, P. R.; BRISOLARA, L. B. B.. Aplicação de Redes de Sensores sem Fio para o Manejo de Pragas em Pomares: avaliação de módulos de comunicação. In: Congresso de Inovação Tecnológica da Universidade Federal de Pelotas, III., 2019. **Anais...** Pelotas, 2019.

HE, K.; ZHANG, X.; REN, S.; SUN, J.. Deep residual learning for image recognition,” In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), **Anais...** 2016, pp. 770–778.

MARTINS, V. A. M.; FREITAS, L. C.; DE AGUIAR, M. S.; BRISOLARA, L. B. B.; FERREIRA, P. R.. Deep Learning applied to the Identification of Fruit Fly in Intelli-gent Traps. In: IX BRAZILIAN SYMPOSIUM ON COMPUTING SYSTEMS ENGINEERING (SBESC), 2019., 2019. **Anais...** [S.l.: s.n.], 2019. p.1–8.

REMBOSKI, T. B. ; DE SOUZA, W. D. ; DE AGUIAR, M. S. ; FERREIRA, P. R.. Identification of fruit fly in intelligent traps using techniques of digital image processing and machine learning. In: 33rd ANNUAL ACM SYMPOSIUM ON APPLIED COMPUTING (SAC), 2018, Pau. **Anais...** , 2018.

XIE, S.; GIRSHICK, R.; DOLLAR, P.; TU, Z.; HE, K.. Aggregated residual transformations for deep neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), **Anais...** 2017, pp.5987–5995.