

RESSALVA

Atendendo solicitação do(a) autor(a) o texto completo desta Tese será disponibilizado somente a partir de 24/01/2024.



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”

Programa de Pós-Graduação em Engenharia Elétrica

**Redes Neurais Profundas com Redução de Pesos para Uso
em Hardwares de Baixo Custo Aplicadas à Identificação de
Espécies Invasoras em Cultivares**

Luiz Carlos Marques Junior

Bauru
2023



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”

Programa de Pós-Graduação em Engenharia Elétrica

**Redes Neurais Profundas com Redução de Pesos para Uso
em Hardwares de Baixo Custo Aplicadas à Identificação de
Espécies Invasoras em Cultivares**

LUIZ CARLOS MARQUES JUNIOR

Tese apresentada à Faculdade de Engenharia de Bauru da Universidade Estadual Paulista “Júlio de Mesquita Filho” para obtenção do título de Doutor em Engenharia Elétrica, pelo Programa de Pós-Graduação em Engenharia Elétrica.

Orientador: Prof. Dr. José Alfredo Covolan
Ulson

Bauru
2023

Marques Junior, Luiz Carlos.

Redes Neurais Profunds com Redução de Pesos
para Uso em Hardwares de Baixo Custo Aplicados à
Identificação de Espécies Invasoras em Cultivares
/Luiz Carlos Marques Junior, 2023

96 f.:il.

Orientador: José Alfredo Covolan Ulson

Tese (Doutorado) - Universidade Estadual
Paulista (Unesp). Faculdade de Engenharia, Bauru,
2023

1. Agroindústria. 2. Plantas Daninhas. 3.
Aprendizado Profundo. 4. Redes Neurais
Convolucionais. 5. Quantização Pós Treinamento.
I. Universidade Estadual Paulista. Faculdade de
Engenharia. II. Título.

ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE LUIZ CARLOS MARQUES JUNIOR, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, DA FACULDADE DE ENGENHARIA - CÂMPUS DE BAURU.

Aos 24 dias do mês de julho do ano de 2023, às 08:30 horas, por meio de Videoconferência, realizou-se a defesa de TESE DE DOUTORADO de LUIZ CARLOS MARQUES JUNIOR, intitulada **REDES NEURAIS PROFUNDAS COM REDUÇÃO DE DIMENSÃO DOS PESOS PARA USO EM HARDWARES DE BAIXO CUSTO APLICADAS À IDENTIFICAÇÃO DE ESPÉCIES INVASORAS EM CULTIVARES**. A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. JOSÉ ALFREDO COVOLAN ULSON (Orientador(a) - Participação Virtual) do(a) Departamento de Engenharia Eletrica / Faculdade de Engenharia de Bauru UNESP, Prof. Dr. DANILO HERNANE SPATTI (Participação Virtual) do(a) Departamento de Sistemas de Computação / Instituto de Ciências Matemáticas e de Computação/USP, Prof. Dr. ADRIANO DE SOUZA MARQUES (Participação Virtual) do(a) Engenharia da Computação / Instituto Federal de São Paulo - IFSP - Câmpus Birigui, Prof. Dr. PEDRO DE OLIVEIRA CONCEIÇÃO JÚNIOR (Participação Virtual) do(a) Departamento de Engenharia Eletrica e de Computacao / Escola de Engenharia de Sao Carlos - Universidade de Sao Paulo, Prof. Dr. MARCELO NICOLETTI FRANCHIN (Participação Virtual) do(a) Departamento de Engenharia Eletrica / Faculdade de Engenharia de Bauru - UNESP. Após a exposição pelo doutorando e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, o discente recebeu o conceito final: aprovado . Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Prof. Dr. JOSÉ ALFREDO COVOLAN ULSON



Epígrafe

*“O conhecimento abre portas que não podem ser fechadas,
rompe barreiras que nunca foram cruzadas e nos faz entender
que na vida somos eternos aprendizes.”*

Agradecimentos

Gostaria de agradecer primeiramente a Deus, por ter me capacitado e permitido a execução desta pesquisa que culminou nesta tese.

A minha esposa pelo apoio, compreensão e companheirismo em todos os momentos.

A família, em especial aos meus pais, por terem me dado a possibilidade e por ter me apoiado nos estudos.

Ao meu orientador Prof. Dr. José Alfredo Covolan Ulson, pelo apoio, tempo dedicado a este trabalho e confiança depositada em mim na elaboração desta pesquisa.

A todos os professores e colegas que passaram em minha vida e gentilmente contribuíram na formação de meu conhecimento.

Resumo

O aumento na demanda por alimentos bem como a alta no preço dos insumos para sua produção são fatores que afetam a agroindústria e a população. Neste cenário, outro desafio é o controle de pragas e doenças, dentre essas pragas destacam-se as espécies invasoras (plantas daninhas), nos cultivares de milho e soja, que são os grãos mais produzidos no mundo. Atualmente, a identificação dessas plantas é um processo insalubre, laborioso e sujeito a falhas quando realizado pelo método manual. Quando o controle é feito por máquinas de maneira não seletiva, o agricultor estará beneficiando espécies resistentes tornando o manejo ineficiente, além de causar impactos ambientais pelo uso excessivo de herbicidas. Portanto, existe a necessidade de um sistema embarcado de baixo custo para a identificação de plantas daninhas que permita seu uso em tratores ou diretamente em plataformas robóticas. Com base nestes desafios, para este trabalho foi proposto e explorado o processo de identificar e localizar espécies invasoras nas culturas de milho e soja por meio de um sistema de inteligência artificial com quantização e conversão de pesos neurais em um *hardware* de baixo custo. Para isso fez-se uso da visão computacional associada ao aprendizado profundo (*Deep Learning*) no treinamento de arquiteturas de redes neurais convolucionais (*Convolutional Neural Network*) *Yolov5* em um conjunto de imagens público das plantas daninhas prejudiciais às culturas de milho e soja. As arquiteturas foram analisadas com base nas métricas matriz de confusão, precisão, revocação, média das precisões médias (*mAP*) com limiar de 0,5 e média das precisões médias com limiar variando 0,5 a 0,95 (*mAP* 0,5:0,95). Após o treinamento fez-se uso da quantização pós treinamento para comprimir os pesos neurais para os formatos inteiro de 8 bits (*int8*), ponto flutuante de 16 bits (*fp16*), e a conversão para *ONNX* (*Open Neural Network Exchange*). Estes pesos juntamente com as imagens de validação foram transferidos para plataforma *Raspberry Pi 4 model B*, onde, realizou-se a avaliação dos modelos com as métricas descritas. Em seguida foi feita a inferência nas imagens do conjunto de validação. Os melhores resultados de validação foram obtidos com a rede *Yolov5l* com quantização em *int8* com *mAP* 0,5 de 0,815, precisão de 0,834, revocação de 0,772 e *mAP* 0,5:0,95 de 0,413, na inferência a rede obteve 1 imagem inferida a cada 5,685 segundos. O modelo com menor tempo de inferência foi a *Yolov5n* com pesos convertidos para *ONNX*, onde foi obtido uma imagem inferida a cada 0,376 segundo. Neste contexto, os resultados apontam que a abordagem proposta é viável para utilização em campo em plataformas tratorizadas/robotizadas. Tem-se como vantagens a possibilidade da aplicação seletiva de defensivos, poupando o meio ambiente, reduzindo custos e aumentando a eficiência na erradicação de plantas daninhas prejudiciais às culturas de milho e soja.

Palavras-chave: Agroindústria, Plantas Daninhas, Aprendizado Profundo, Redes Neurais Convolucionais, Quantização Pós Treinamento.

Abstract

The increase in the demand for food as well as the rise in the price of the inputs for its production are factors that affect the agro-industry and the population. In this scenario, another challenge is the control of pests and diseases, including invasive species (weeds) in corn and soybean cultivars, which are the most produced grains in the world. Currently, identifying these plants is an unhealthy, laborious and error-prone process when carried out manually. When control is carried out by machines in a non-selective way, the farmer will be benefiting resistant species, making management inefficient and causing environmental impacts through the excessive use of herbicides. Therefore, there is a need for a low-cost embedded system for weed identification that can be used on tractors or directly on robotic platforms. Based on these challenges, this work proposed and explored the process of identifying and locating invasive plant species in corn and soybean crops using an artificial intelligence system with quantization and conversion of neural weights on low-cost hardware. For this purpose, computer vision associated with deep learning was used to train Yolov5 convolutional neural network architectures on a set of public images of weeds harmful to corn and soybean crops. The architectures were analyzed based on the following metrics: confusion matrix, precision, recall, mean average precision (mAP) with a threshold of 0.5 and mean average precision with a threshold ranging from 0.5 to 0.95 (mAP 0.5:0.95). After training, post-training quantization was used to compress the neural weights into 8-bit integer (int8) and 16-bit floating point (fp16) formats, and conversion to ONNX (Open Neural Network Exchange). These weights and the validation images were transferred to the Raspberry Pi 4 model B platform, where the models were evaluated using the metrics described. Inference was then made on the images in the validation set. The best validation results were obtained with the Yolov5l network with int8 quantization with mAP 0.5 of 0.815, precision of 0.834, recall of 0.772 and mAP 0.5:0.95 of 0.413. In inference, the network obtained 1 inferred image every 5.685 seconds. The model with the fastest inference time was Yolov5n with weights converted to ONNX, where an inferred image was obtained every 0.376 seconds. In this context, the results show that the proposed approach is viable for use in the field on tractor/robot platforms. The advantages include the possibility of selective application of defensives, which conserves the environment, reduces costs and increases efficiency in eradicating weeds that are harmful to corn and soybean crops.

Keywords: Agroindustry, Weeds, Deep Learning, Convolutional Neural Networks, Post-Training Quantization.

Lista de Figuras

Figura 1 - Produção de Soja no Brasil.....	23
Figura 2 - Produção de Milho no Brasil.....	23
Figura 3 - Técnicas Tradicionais para o Controle de Plantas Daninhas (a) Controle Mecânico, (b) Controle Cultural, (c) Controle Biológico, (d) Controle Químico	26
Figura 4 - Representação de Imagem em Tons de Cinza (a) Figura Pixelizada, (b) Figura Pixelizada com o Valores do Pixels, (c) Valores dos Pixels	27
Figura 5 - Representação de Imagem em Colorida e suas Dimensões.....	28
Figura 6 - Representação do Neurônio Biológico (a) Neurônio Matemático de McCulloch-Pitts (b)	31
Figura 7 - Estrutura de uma Rede Neural Tradicional e uma Rede Profunda.....	32
Figura 8 - Rede Neural com Retro propagação e Otimização via Gradiente Descendente	33
Figura 9 - Rede Neural Convolucional proposta por Lecun, Y	34
Figura 10 - Classificação Binária e de Múltiplas Classes	35
Figura 11 - Classificação, Classificação com Localização e Detecção de Objetos	36
Figura 12 - Imagem de Entrada e Máscara de Entrada para Segmentação Semântica.	37
Figura 13 - Representação de Dados em FP16 e FP32	38
Figura 14 - Representação de Dados em FP32, FP16 e INT8.....	38
Figura 15 - Conversão de Pesos Neurais para ONNX e Implantação em Diferentes Dispositivos	39
Figura 16 - Curva Precisão-Revocação	42
Figura 17 - Representação de Intersecção sob União de (a) 0,2, (b) 0,5 e (c) 0,9.....	43
Figura 18 - Matriz de Confusão Indicando a Classificação Entre Plantas de Folha Larga e Folha Estreita.....	44
Figura 19 - Imagens Segmentadas Utilizadas para Classificação de Plantas Daninhas	45
Figura 20 - Plataforma de Coleta de Imagens BoniRob e Plantas Analisadas	47
Figura 21 - Aparelho WeedLogger para coleta de imagens com coordenadas GPS	47
Figura 22 - Fluxo de Procedimentos Metodológicos da Tese.....	51
Figura 23 - Espécies de plantas daninhas do 4WeedDataset: (a) Carrapicho de Carneiro, (b) Rabo de Raposa, (c) Caruru Gigante e (d) Erva de Santiago Gigante	52
Figura 24 - Formato de Anotação Yolo.....	53
Figura 25 - Arquitetura Yolov5 composta por: Espinha Dorsal (Backbone) CSPDarknet, Pescoço (Neck) PANet e Cabeça (Head) Yolo	55
Figura 26 - Hiperparâmetros de Treinamento e Augmentação para as Redes Yolov5 .	56

Figura 27 - Trecho de código para Quantização nos Formatos fp16 e int8	59
Figura 28 - Trecho de código para Conversão no Formato ONNX	60
Figura 29 - Principais Componentes do Módulo Raspberry Pi 4 Model B 4GB	61
Figura 30 - Placa Raspberry Pi Model B 4GB Utilizada nos Experimentos	61
Figura 31 - Matriz de Confusão Yolov5n com pesos em fp32	64
Figura 32 - Matriz de Confusão Yolov5s com pesos em fp32	65
Figura 33 - Matriz de Confusão Yolov5m com pesos em fp32	65
Figura 34 - Matriz de Confusão Yolov5l com pesos em fp32	66
Figura 35 - Matriz de Confusão Yolov5x com pesos em fp32	66
Figura 36 - Matriz de Confusão Yolov5n com pesos quantizados fp16	70
Figura 37 - Matriz de Confusão Yolov5n com pesos quantizados int8	70
Figura 38 - Matriz de Confusão Yolov5n com pesos convertidos para ONNX	71
Figura 39 - Matriz de Confusão Yolov5s com pesos quantizados fp16	74
Figura 40 - Matriz de Confusão Yolov5s com pesos quantizados int8	74
Figura 41 - Matriz de Confusão Yolov5s com pesos convertidos para ONNX	75
Figura 42 - Matriz de Confusão Yolov5m com pesos quantizados fp16	77
Figura 43 - Matriz de Confusão Yolov5m com pesos quantizados int8	78
Figura 44 - Matriz de Confusão Yolov5m com pesos convertidos para ONNX	78
Figura 45 - Matriz de Confusão Yolov5l com pesos quantizados fp16	81
Figura 46 - Matriz de Confusão Yolov5l com pesos quantizados int8	81
Figura 47 - Matriz de Confusão Yolov5l com pesos convertidos para ONNX	82
Figura 48 - Matriz de Confusão Yolov5x com pesos quantizados fp16	84
Figura 49 - Matriz de Confusão Yolov5x com pesos quantizados int8	85
Figura 50 - Matriz de Confusão Yolov5x com pesos convertidos para ONNX	85

Lista de Tabelas

Tabela 1 - Distribuição de Imagens por Espécie e Rótulos de Plantas Daninhas.....	52
Tabela 2 - Variações de modelos da Rede Neural Yolov5 e suas características	55
Tabela 3 - Número de Imagens por Espécie de Planta no Dataset de Treinamento	57
Tabela 4 - Número de imagens por Espécie de Planta no Dataset de Validação	57
Tabela 5 - Resultados de Validação Redes Yolov5 pesos em fp32.....	68
Tabela 6 - Pesos Neurais Rede Yolov5n Quantizados e Convertidos.....	69
Tabela 7 - Resultados de Validação Rede Yolov5n pesos Quantizados e Convertidos	72
Tabela 8 - Pesos Neurais Rede Yolov5s Quantizados e Convertidos	73
Tabela 9 - Resultados de Validação Rede Yolov5s pesos Quantizados e Convertidos	76
Tabela 10 - Pesos Neurais Rede Yolov5m Quantizados e Convertidos.....	76
Tabela 11 - Resultados de Validação Rede Yolov5m pesos Quantizados e Convertidos	79
Tabela 12 - Pesos Neurais Rede Yolov5m Quantizados e Convertidos.....	80
Tabela 13 - Resultados de Validação Rede Yolov5l pesos Quantizados e Convertidos	83
Tabela 14 - Pesos Neurais Rede Yolov5x Quantizados e Convertidos	83
Tabela 15 - Resultados de Validação Rede Yolov5l pesos Quantizados e Convertidos	87
Tabela 16 - Resultados de Inferência Rede Yolov5n pesos Quantizados e Convertidos	88
Tabela 17 - Resultados de Inferência Rede Yolov5s pesos Quantizados e Convertidos	88
Tabela 18 - Resultados de Inferência Rede Yolov5m pesos Quantizados e Convertidos	89
Tabela 19 - Resultados de Inferência Rede Yolov5l pesos Quantizados e Convertidos	89
Tabela 20 - Resultados de Inferência Rede Yolov5x pesos Quantizados e Convertidos	90
Tabela 21 - Resultados de Inferência Redes Yolov5 com pesos Quantizados e Convertidos.....	90

Lista de Abreviaturas e Siglas

AP	Average Precision
CEPEA	Centro de Estudos Avançados em Economia Aplicada
CNN	Convolutional Neural Network
COCO	Common Objects in Context
CPU	Central Processing Unit
CSPNET	Cross Stage Partial Network
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
ESALQ	Escola Superior Luiz de Queiroz
FAO	Food and Agriculture Organization of the United Nations
FAOSTAT	Food and Agriculture Statistics
FLOPS	Floating Point Per Second
FP16	Float Point 16
FP32	Float Point 32
FPS	Frames por Segundo
GPU	Graphics Processing Units
INT8	Integer 8 bits
IOU	Intersection Over Union
mAP	Mean Average Precision
MB	Megabyte
PANET	Path Aggregation Network
PIB	Produto Interno Bruto
RNA	Redes Neurais Artificiais
RNC	Redes Neurais Convolucionais
SLIC	Simple Linear Iterative Clustering
USDA	United States Department of Agriculture
VANT	Veículo Aéreo Não Tripulado

Sumário

CAPÍTULO 01: INTRODUÇÃO	17
1.1. Motivação.....	17
1.2. Objetivo Geral.....	19
1.3. Objetivos Específicos.....	19
1.4. Estrutura da Tese	20
CAPÍTULO 2: REVISÃO BIBLIOGRÁFICA	22
2.1. Produção Agrícola no Brasil e no Mundo.....	22
2.2. Plantas Daninhas	24
2.2.1. Técnicas Tradicionais para o Manejo de Plantas Daninhas	25
2.3. Visão Computacional	26
2.4. Aprendizado de Máquina	28
2.5. Redes Neurais Artificiais	30
2.6. Aprendizado Profundo	32
2.7. Redes Neurais Convolucionais	34
2.8. Classificação, Detecção de Objetos e Segmentação Semântica em Imagens 35	
2.8.1. Classificação.....	35
2.8.2. Detecção de Objetos	36
2.8.3. Segmentação Semântica.....	36
2.9. Quantização de Pesos Neurais e Conversão para ONNX.....	37
2.9.1. Quantização para Ponto Flutuante de 16 Bits	37
2.9.2. Quantização para Inteiro de 8 Bits	38
2.9.3. Conversão para ONNX	39
2.10. Métricas para Avaliação de Resultados.....	40
2.11. Abordagens baseadas em Redes Neurais Profundas para a Identificação de Plantas Daninhas	45
CAPÍTULO 3: METODOLOGIA	50
3.1. Procedimentos Metodológicos	50
3.2. Conjunto de Imagens.....	51
3.3. Algoritmo Yolov5.....	53
3.4. Treinamento Modelos Yolov5.....	57
3.5. Quantização e Conversão dos Pesos Neurais	58
3.6. Hardware de Inferência.....	60

CAPÍTULO 4: RESULTADOS E DISCUSSÕES	63
4.1. Resultado Treinamento Modelos <i>Yolov5</i>	63
4.2. Resultados Quantização e Conversão <i>Yolov5n</i>	69
4.3. Resultados Quantização e Conversão <i>Yolov5s</i>	72
4.4. Resultados Quantização e Conversão <i>Yolov5m</i>	76
4.5. Resultados Quantização e Conversão <i>Yolov5l</i>	80
4.6. Resultados Quantização e Conversão <i>Yolov5x</i>	83
4.7. Resultados de Inferência Redes <i>Yolov5</i> com Pesos Quantizados e Convertidos.....	87
CAPÍTULO 5: CONCLUSÕES E TRABALHOS FUTUROS.....	93
5.1. Conclusões.....	93
5.2. Propostas para Trabalhos Futuros	93
REFERÊNCIAS	94

CAPÍTULO 01: INTRODUÇÃO

Neste capítulo é apresentado o contexto e relevância deste projeto de pesquisa, bem como a motivação, os objetivos gerais, específicos e a estrutura da tese.

1.1. Motivação

De acordo com o relatório da FAO é estimado que a população mundial atinja a marca de 9,7 bilhões de habitantes em 2050, consequentemente aumentando a demanda por alimentos. É estimado também que cerca de dois terços da população mundial viverá em áreas urbanas e um terço em áreas rurais (FAO, 2017).

Nesse contexto, o Brasil ocupa papel de destaque na agroindústria mundial, sendo o maior exportador de carne bovina do mundo e o segundo maior em exportação de grãos, de acordo com a EMBRAPA (EMBRAPA, 2021). Entretanto, um fator limitante para a produtividade agrícola é o controle de pragas e doenças como insetos e plantas daninhas, que podem causar perdas significativas. Segundo a EMBRAPA as plantas daninhas causam perdas diretas na produção via competição por recursos (sol, nutrientes do solo) e alelopatia (efeito que uma planta exerce sobre a outra via a liberação de compostos químicos) reduzindo a capacidade de produção das plantas, assim como perdas indiretas, como o aumento do custo de produção, dificuldade de colheita devido à infestação de plantas daninhas e depreciação na qualidade do alimento produzido. Caso não seja feito manejo das plantas daninhas as perdas podem chegar a 90% e com manejo atualmente as perdas são estimadas entre 13 e 15% para a produção de grãos (EMBRAPA, s.d).

Com base nesse cenário, a identificação de plantas daninhas é essencial para manter a produtividade alta, os custos de produção baixos e evitar o uso em excesso de herbicidas, prejudicando o meio ambiente. Porém, a identificação de pragas e doenças pelo método manual em especial as plantas daninhas é um trabalho sujeito a falhas humanas devido à grande extensão das plantações, também por ser um processo repetitivo e laborioso. Caso feito esse controle por máquinas de maneira não seletiva, o agricultor estará beneficiando espécies resistentes o que tornará esse manejo ineficiente.

Portanto, automatizar o processo de identificação de diferentes espécies de plantas daninhas para a aplicação localizada de defensivos agrícolas ou remoção via outros métodos é primordial. Nos últimos anos uma solução promissora para este desafio tem surgido que é a visão computacional associada com algoritmos de aprendizado de máquina (*machine learning*) e recentemente o aprendizado profundo (*deep learning*) bem como o uso de redes neurais convolucionais (*convolutional neural networks*).

Na literatura alguns trabalhos destacam-se relacionados a esse tema, como o uso de redes neurais convolucionais (CNNs) para a classificação de imagens de plantas daninhas (DOS SANTOS FERREIRA et al., 2017; SUH et al., 2018; TEIMOURI et al., 2018; TRONG et al., 2020) detecção de plantas daninhas em tempo real utilizando plataformas robotizadas (MILIOTO et al., 2018; PARTEL et al., 2019). Em um trabalho anterior (MARQUES JUNIOR; ULSON, 2018) foi proposto a comparação entre modelos de aprendizado profundo para a classificação de imagens, distinguindo solo, soja, plantas daninhas de folha larga e folha estreita, neste trabalho o melhor resultado de precisão foi obtido com a rede VGG16 com 96,5%. Entretanto, percebeu-se que além de classificar as plantas daninhas era necessário detectá-las, diferenciando-as do plano de fundo, assim obtendo a localização das plantas e embarcando os algoritmos de inteligência artificial em plataformas robóticas ou tratorizadas, criando um sistema de aplicação automática de defensivos com aplicação localizada e seletiva.

Foi proposto então em outro trabalho (MARQUES JUNIOR; ULSON, 2021) a implementação e validação dos modelos de rede *Yolov5* (ULTRALYTICS, 2020) em um conjunto de imagens (*dataset*) para identificação e detecção de plantas daninhas resistentes ao herbicida Glifosato. Neste trabalho o melhor tempo de inferência (imagens analisadas pelo algoritmo de IA) foi de 62 FPS (frames por segundo) com a rede *Yolov5s*, já a melhor média das precisões médias com 0,5 de limiar (*mAP 0,5*) foi obtido com a rede *Yolov5m* com *mAP 0,5* de 0,77 e 26 FPS, inferindo localmente em um notebook *Samsung Odyssey* com processador Core i7, 8 gigas de memória ram e placa gráfica Nvidia GTX 1050 com 4 gigas.

Apesar de promissor, o custo de implementação de um sistema como o descrito, tanto para a implantação em larga escala quanto para poucas máquinas, seria alto e inviável para os agricultores. Tendo em vista essa limitação e desafio proposto para

este trabalho foi explorar a implementação de redes neurais profundas dedicadas a detecção de objetos em um *hardware* de baixo custo e processamento.

Portanto, neste trabalho os distintos modelos da rede neural *Yolov5* foram treinados em um *dataset* público de plantas daninhas prejudiciais às culturas de milho e soja. Posteriormente, foi feita a quantização dos pesos neurais para as representações de ponto flutuante de 16 bits (*fp16*) inteiro de 8 bits (*int8*) e conversão desses pesos para o formato ONNX (*Open Neural Network Exchange*). Para a quantização os pesos neurais são convertidos do formato *pytorch* para *tensorflow* em seguida para as representações *tflite.fp16* e *tflite.int8*. Já a conversão para ONNX é feita por meio da função *torch.onnx.export* nativa do *pytorch*.

De posse dos pesos neurais, avaliações de desempenho foram realizadas a fim de constatar a viabilidade da abordagem proposta. Nos experimentos foi avaliada a inferência e identificação de plantas daninhas com os dados dos pesos neurais no módulo *Raspberry Pi 4 model B* de 4GB. De acordo com os nossos conhecimentos não existem trabalhos que exploraram a abordagem de quantização e conversão de pesos neurais para implantar um sistema de baixo custo aplicado na identificação de plantas daninhas prejudiciais nas culturas de milho e soja.

1.2. Objetivo Geral

Implementar uma abordagem baseada em redes neurais artificiais para identificação de plantas daninhas nas culturas de milho e soja a partir de imagens no espectro visível, possibilitando que esse sistema seja embarcado em veículos tratorizados.

1.3. Objetivos Específicos

Tem-se como objetivos específicos para este trabalho:

- Treinar os distintos modelos da *Yolov5* em um conjunto de imagens públicas de plantas daninhas prejudiciais as culturas de milho e soja, sendo essas plantas: Carrapicho de Carneiro, Rabo de Raposa, Caruru Gigante e Erva de Santiago Gigante;

- Gerar as matrizes de confusão e demais métricas para avaliação de assertividade desses modelos;
- Quantizar e converter os pesos neurais para os formatos *int8*, *fp16* e *ONNX* e transferir os arquivos de pesos para o modulo *Raspberry Pi 4 model B* de 4GB.
- No modulo *Raspberry* gerar as métricas descritas no item anterior para avaliar assertividade, bem como avaliações de tempo de inferência por imagem para cada peso neural.
- De posse dos resultados de assertividade e inferência discutir os resultados e apontar quais modelos apresentaram melhor desempenho.

1.4. Estrutura da Tese

Os próximos capítulos desta tese estão organizados da seguinte maneira:

O capítulo 2 apresenta a revisão bibliográfica, contendo uma contextualização mais aprofundada a respeito da produção alimentícia levando em conta o Brasil e o mundo, posteriormente é dado enfoque no desafio principal deste trabalho que é o controle de plantas daninhas e quais são técnicas utilizadas. Em seguida os são explanados os conceitos, definições e abordagens no uso de *machine learning*, redes neurais artificiais e visão computacional nas tarefas de classificação, detecção de objetos e segmentação semântica. Em sequência as métricas para avaliação dos resultados dessas abordagens são descritos. Por fim o processo de quantização e conversão de pesos neurais é descrito seguido pelos principais trabalhos relacionados ao escopo da tese.

No capítulo 3 inicialmente é apresentado o fluxograma com os procedimentos metodológicos da tese, em seguida é descrita a metodologia utilizada desde a preparação do conjunto de imagens, arquitetura e características da rede neural *Yolov5*, o ambiente e dependências utilizados para o treinamento dos modelos, por fim é descrito o processo de quantização e conversão dos pesos neurais, bem como apresentado *hardware* explorado nos experimentos e seus principais componentes.

No capítulo 4 são apresentados os resultados e discussões dos procedimentos experimentais, inicialmente são avaliados os resultados de matrizes de confusão, precisão, revocação média das precisões médias com limiar de 0,5 (mAP 0,5) e com limiar variando de 0,5 à 0,95 (mAP 0,5: 0,95) na representação de ponto flutuante e 32 *bits*. Em seguida as mesmas métricas apontadas são avaliadas na *Raspberry Pi 4 model B* de 4GB, com os pesos neurais de cada rede quantizado e convertido. Por fim é feito o levantamento de tempo de inferência de imagens para cada modelo na *Raspberry Pi 4 model B* de 4GB.

Ao final o capítulo 5 traz as conclusões, contribuições obtidas com este trabalho bem como perspectivas de melhoras para trabalhos futuros.

CAPÍTULO 5: CONCLUSÕES E TRABALHOS FUTUROS

Neste capítulo são apresentadas as conclusões com base nos resultados obtidos, também são elencadas as principais contribuições dessa tese e propostas abordagens para trabalhos futuros.

5.1. Conclusões

Para este trabalho foi proposto um sistema para identificar as espécies invasoras para as culturas de milho e soja utilizando as arquiteturas *Yolov5* quantizadas para os formatos fp16, int8 e convertidas para *ONNX* na placa *Raspberry Pi model B 4GB*.

Como principal contribuição deste trabalho, foi proposto e demonstrado uma abordagem para implantar um sistema de inteligência artificial com quantização e conversão de pesos aplicados na identificação e detecção seletiva de plantas daninhas prejudiciais as culturas de milho e soja em um *hardware* de baixo custo. Tal abordagem traz como vantagens: i) Permite a identificação *on-the-go* de plantas daninhas em plataformas tratorizadas; ii) Permite a erradicação de plantas daninhas de forma seletiva; iii) Promove a economia de insumos agrícolas; iv) Contribui para a redução de impacto ambiental; v) O sistema apresenta baixo custo para implementação. Como principal desvantagem pode-se citar: i) A velocidade de deslocamento em campo é sensível ao desempenho do *hardware* de baixo custo utilizado.

5.2. Propostas para Trabalhos Futuros

Como sugestões para trabalhos futuros sugere-se: i) Criação de vídeos obtidos com câmeras acopladas em sistemas tratorizados para avaliar a inferência das diferentes representações de pesos neurais; ii) Implantação do sistema de identificação de plantas daninhas em uma plataforma robotizada ou acoplado em um trator agrícola para visando apurar o aprendizado para outras espécies de plantas daninhas encontradas no Brasil, iii) Avaliar o desempenho e a velocidade de deslocamento em campo para outras plataformas de hardware de baixo custo.

REFERÊNCIAS

- A. PASZKE et al. **PyTorch: An Imperative Style, High-Performance Deep Learning Library**. In *NeurIPS*, Vancouver, v. 32, p. 8024—8035, 2019.
- AGGARWAL, V. et al. **4Weed Dataset: Annotated Imagery Weeds Dataset**. Cite arXiv:2204.00080, 2022.
- AI. STANFORD. **Tutorial 1: Image Filtering**. Estados Unidos 2022. Disponível em: <https://ai.stanford.edu/~syueung/cvweb/tutorial1.html> . Acesso em 25/05/2022.
- AHMED. S et al. OpTorch: **Optimized deep learning architectures for resource limited environments**. Cite arXiv:2105.00619, 2021.
- ALLABOUTCIRCUITS. Neural Network Quantization: **What Is It and How Does It Relate to TinyML**. Estados Unidos, 2022. Disponível em: <https://www.allaboutcircuits.com/technical-articles/neural-network-quantization-what-is-it-and-how-does-it-relate-to-tiny-machine-learning/> . Acesso em 30/04/2022.
- BISONG. E. Google Colaboratory. In: **Building Machine Learning and Deep Learning Models on Google Cloud Platform**. Apress, Berkeley, CA, 2019.
- BOCHKOVSKIY. A et al. **YOLOv4: Optimal Speed and Accuracy of Object Detection**. Cite arXiv:2004.10934, 2020.
- CEPEA. **PIB-AGRO/CEPEA**. Brasil, 2022. Disponível em: <https://www.cepea.esalq.usp.br/br/releases/pib-agro-cepea-pib-do-agro-cresce-8-36-em-2021-participacao-no-pib-brasileiro-chega-a-27-4.aspx>. Acesso em 07/05/2022.
- DECI.AI. **The Ultimate Guide to Dee Learning Model Quantization and Quantization-Aware Training**. Estados Unidos, 2023. Disponível em: <https://deci.ai/quantization-and-quantization-aware-training/>. Acesso em 22/03/2023.
- DEEPAI. **Convolutional Neural Network**. Estados Unidos, 2022. Disponível em: <https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network>/>. Acesso em 22/04/2022.
- DEEP LEARNING BOOK. **Capítulo 3 – O Que São Redes Neurais Artificiais Profundas ou Deep Learning, Biológico e Matemático**. Estados Unidos, 2022. Disponível em: <https://www.deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/> Acesso em 21/04/2022.
- DEEP LEARNING BOOK. **Capítulo 4 – O Neurônio, Biológico e Matemático**. Estados Unidos, 2022. Disponível em: <http://deeplearningbook.com.br/o-neuronio-biologico-e-matematico/>. Acesso em 21/04/2022.
- DENG, J. et al. **ImageNet: A Large-Scale Hierarchical Image Database**. In *CVPR.IEEE* , Miami, v. 1, p. 248 - 255, 2009.

DOS SANTOS FERREIRA, A. et al. **Weed detection in soybean crops using convnets**. *Computers and Electronics in Agriculture*, Elsevier, v. 143, p. 314-324, 2016.

EMBRAPA. **Plantas Daninhas**. Brasil, s.d. Disponível em: <https://www.embrapa.br/tema-plantas-daninhas/sobre-o-tema>. Acesso em 12/10/2021.

EMBRAPA. **Estudos Socioeconômicos e Ambientais**. Brasil, 2021. Disponível em <https://www.embrapa.br/busca-de-noticias/-/noticia/62619259/brasil-e-o-quarto-maior-produtor-de-graos-e-o-maior-exportador-de-carne-bovina-do-mundo-diz-estudo>. Acesso em 08/10/2021.

EVERINGHAM, M. et al. **The pascal visual object classes challenge: A retrospective**. *International Journal of Computer. Vision*, v. 111, no. 1, p. 98–136, 2015.

FELIPEFLOP. **Raspberry Pi 4 Model B Anatel**. Brasil, 2022. Disponível em: <https://www.filipeflop.com/produto/raspberry-pi-4-model-b/>. Acesso em 22/04/2022.

FAO. **The future of food and agriculture – Trends and challenges**. Roma, 2017. Roma, 2017. Disponível em: <https://www.fao.org/3/i6583e/i6583e.pdf>. Acesso em 15/07/2021.

FAO. **Agricultural production statistics. 2000–2020. FAOSTAT Analytical Brief Series No. 41**. Roma, 2022. Disponível em: <https://www.fao.org/3/cb9180en/cb9180en.pdf> . Acesso em 10/05/2022.

FORAGES. **Describe the five general categories of weed control methods**. Estados Unidos, 2021. Disponível em: <https://forages.oregonstate.edu/nfgc/eo/onlineforagecurriculum/instructormaterials/availabletopics/weeds/control>. Acesso em 15/05/2022.

GROW. **Mechanical Weed Control**. Estados Unidos, 2021. Disponível em: <https://growiwm.org/mechanical-weed-control/>. Acesso em 15/05/2022.

HE, K. et al. **Deep Residual Learning for Image Recognition**. In *CVPR.IEEE*, Las Vegas, v. 1, p. 770-778, 2016.

HOWARD, A. G. et al. **MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications**. Cite arxiv:1704.04861, 2017.

KRIZHEVSKY, A. et al. **ImageNet Classification with Deep Convolutional Neural Networks**. *Advances in Neural Information Processing Systems*, San Diego, v. 25, p. 1097-1105, 2012.

LECUN, Y. et al. **Gradient-based learning applied to document recognition**. *Proceedings of the IEEE*, v. 88, p. 2278–2324, 1998.

MARQUES JUNIOR, L. C.; ULSON, J. A. C. **Utilização De Redes Neurais Convolucionais Em Processos De Classificação De Espécies Vegetais Na Agroindústria Da Soja**. In: *INDUSCON. IEEE*, São Paulo, v. 1, p. 2226-2231, 2018.

MARQUES JUNIOR, L. C.; ULSON, J. A. C. **Real Time Weed detection using Computer vision and Deep learning**. In: *INDUSCON. IEEE*, São Paulo, v. 1, p. 1131-1137, 2021.

MEDIUM. **Hardware for Deep Learning**. Part 3: GPU. Estados Unidos, 2018. Disponível em: <<https://blog.inten.to/hardware-for-deep-learning-part-3-gpu-8906c1644664#ea08>> Acesso em 22/04/2022.

MILIOTO, A. et al. **Real-Time Semantic Segmentation of Crop and Weed for Precision Agriculture Robots Leveraging Background Knowledge in CNNs**. In: *ICRA. IEEE*, Áustria, v. 1, p. 2229–2235, 2018.

MIT. **Machine learning, explained**. Estados Unidos, 2021. Disponível em: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>>. Acesso em 21/04/2022.

MIT. **Robotic Manipulation**. Estados Unidos, 2021. Disponível em:<<https://manipulation.csail.mit.edu/segmentation.html>>. Acesso em 23/03/2023.

NSW DPI. **About Weed Biological Control**. Australia, 2022. Disponível em: <https://www.dpi.nsw.gov.au/biosecurity/weeds/weed-control/biological-control/about-weed-biological-control>. Acesso em: 15/05/2022.

OLSEN, A. et al. **DeepWeeds: A Multiclass Weed Species Image Dataset for Deep Learning**. *Sci Rep* 9, 2058 (2019).

PARTEL, V. et al. **Smart sprayer for precision weed control using artificial intelligence: Comparison of deep learning frameworks**. *Association for the Advancement of Artificial Intelligence*. 2019.

PAPERSPACE. **Evaluating Deep Learning Models: The Confusion Matrix, Accuracy, Precision, and Recall**. Estados Unidos, 2020. Disponível em: <https://blog.paperspace.com/deep-learning-metrics-precision-recall-accuracy/>. Acesso em 30/04/2022.

PAPERSPACE. **Evaluating Object Detection Models Using Mean Average Precision (mAP)**. Estados Unidos, 2020. Disponível em: <https://blog.paperspace.com/mean-average-precision/>. Acesso em 30/04/2022.

RAMO, K. **Hands-On Java Deep Learning for Computer Vision**. Birmingham, UK: Packt Publishing, 2019.

REDMON, J. et al. **You Only Look Once: Unified, Real-Time Object Detection**. In *CVPR.IEEE*, Las Vegas, v. 1, p. 779 - 788, 2016.

REDMON, J., FARHADI, A. **YOLO9000: Better, Faster, Stronger**. In *CVPR. IEEE*, p. 6517-6525, 2017.

REDMON, J., FARHADI, A. **Yolov3: An incremental improvement**. Cite arXiv:1804.02767, 2018.

REDMON, J. **Darknet13**. Estados Unidos, 2013. Disponível em: <https://pjreddie.com/darknet/>. Acesso em: 18/03/2022.

REN, S. et al. **Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks**. In *NIPS*, pp. 91–99, 2015.

REN, S. et al. **Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks**. In *Transactions on Pattern Analysis and Machine Intelligence*. IEEE, v.39, p. 1137 - 1149, 2016.

SERMANET, P et al. **Overfeat: Integrated recognition, localization and detection using convolutional networks**. *International Conference on Learning Representations*, Banff, 2014.

SIMONYAN, K; ZISSERMAN, A. **Very deep convolutional networks for Large-Scale image Recognition**. *International Conference on Learning Representations*, San Diego, 2015.

SZEGEDY, C. et al. **Going Deeper with Convolutions**. In *CVPR.IEEE*, Boston, v. 1, p. 1 - 9, 2015.

SZEGEDY, C. et al. **Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning**. In *ICLR*, San Juan, 2016.

SZEGEDY, C. et al. **Rethinking the Inception Architecture for Computer Vision**. In *CVPR. IEEE*, Las Vegas, v. 1, p.2818 - 2826, 2016.

SUH, H. K. et al. **Transfer learning for the classication of sugar beet and volunteer potato under eld conditions**. *Biosystems engineering*, Elsevier, v. 174, p. 50-65, 2018.

TEIMOURI, N. et al. **Weed Growth Stage Estimator Using Deep Convolutional Neural Networks**. *Sensors*, vol. 18, no. 5, p. 1580, 2018.

THEENGINEERINGPROJECTS. **What is a Raspberry Pi 4? Pinout, Spec, Projects & Datasheet**. Estados Unidos, 2021. Disponível em: <https://www.theengineeringprojects.com/2021/03/what-is-raspberry-pi-4-pinout-specs-projects-datasheet.html> . Acesso em 13/03/2023

TRONG, V. H. **Late fusion of multimodal deep neural networks for weeds classication**. *Computers and Electronics in Agriculture*, vol. 175, p. 105506, 2020.

ULTRALYTICS. **Yolov5 by Ultralytics**. Estados Unidos, 2020. Disponível em: <https://github.com/ultralytics/yolov5> . Acesso em 15/10/2021.

USDA. **Soybeam Explorer**. Disponível em: <https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=2222000> . Acesso em 15/04/2023.

USDA. **Corn Explorer**. Disponível em: <https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=0440000> . Acesso em 15/04/2023.

W, S. MCCULLOCH.; Pitts, W. **A logical calculus of the ideas immanent in nervous activity**. *The bulletin of mathematical biophysics*, v. 5, p.115–133, 1943.

WANG, K. et al. Panet: **Few shot image semantic segmentation with prototype alignment**. *In ICCV.IEEE*, Seoul, Coreia, p. 9197– 9206, 2019.

WANG, C.Y. et al. CSPNet: **A New Backbone that can Enhance Learning Capability of CNN**. *In CVPR. IEEE*, Washington, DC, p. 1571–1580, 2020.

WEISS, K. et al. A survey of transfer learning. *Journal of Big Data*, vol. 3, p. 9, 2016.

XU, R. et al. **A Forest Fire Detection System Based on Ensemble Learning**. *Forests*, v. 12, no. 2, p. 217, 2021.

ZOPH, B. et al. **Learning Transferable Architectures for Scalable Image Recognition**. Cite arxiv:1707.07012, 2017.