

15º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2024

DEMARCAÇÃO AUTOMATIZADA DE PADRÕES EM TALHÕES AGRÍCOLAS: APLICAÇÃO DE INTELIGÊNCIA ARTIFICIAL E PROCESSAMENTO DE IMAGENS

A. L. S. CONDE¹, H. G. SANTOS², F. S. SANTOS³

¹ Graduando em Engenharia da Computação, IFSP, Campus Birigui, andre.conde@aluno.ifsp.edu.br

² Graduando em Engenharia da Computação, IFSP, Campus Birigui, higor.santos@aluno.ifsp.edu.br

³ Professor Doutor do IFSP, Campus Birigui, fsergio.santos@ifsp.edu.br

Área de conhecimento (Tabela CNPq): 1.03.03.04-9 Sistemas de Informação

RESUMO: A detecção de carregadores em campos de cultivo é um desafio significativo na agricultura de precisão, exigindo tempo e mão de obra qualificada, o que se configura como um obstáculo à plena automatização do setor. A automação desse processo tem o potencial de aumentar a produtividade agrícola e aperfeiçoar a gestão de recursos, sendo um objetivo estratégico para a inovação no campo. Este trabalho propõe o uso de Redes Neurais Convolucionais e Processamento Digital de Imagens para automatizar a detecção de carregadores e a conversão das coordenadas espaciais através de um algoritmo específico. Os testes realizados demonstraram que a solução proposta consegue realizar a detecção de maneira autônoma e com maior rapidez em comparação aos métodos tradicionais. No entanto, a eficácia do sistema é restringida pela técnica de detecção de linhas, especificamente a Transformada de Hough, que ainda apresenta limitações. A metodologia apresenta um potencial promissor, mas sua eficiência está diretamente atrelada ao aprimoramento das técnicas de detecção de linhas, o que indica a necessidade de avanços futuros para superar essas limitações e melhorar a precisão e a confiabilidade do processo.

PALAVRAS-CHAVE: transformada de distancia; fechamento; esqueletização; redes neurais convolucionais; transformada de hough; conversão de coordenadas;

AUTOMATED DEMARCATION OF PATTERNS IN AGRICULTURAL PLANTS: APPLICATION OF ARTIFICIAL INTELLIGENCE AND IMAGE

ABSTRACT: The detection of carriers in agricultural fields is a significant challenge in precision farming, requiring time and skilled labor, which poses an obstacle to the full automation of the sector. Automating this process has the potential to increase agricultural productivity and improve resource management, making it a strategic goal for innovation in the field. This work proposes the use of Convolutional Neural Networks and Digital Image Processing to automate carrier detection and convert spatial coordinates through a specific algorithm. The tests conducted demonstrated that the proposed solution can perform detection autonomously and more quickly compared to traditional methods. However, the system's effectiveness is limited by the line detection technique, specifically the Hough Transform, which still presents limitations. The methodology shows promising potential, but its efficiency is directly tied to the improvement of line detection techniques, indicating the need for future advancements to overcome these limitations and enhance the precision and reliability of the process.

KEYWORDS: distance transform; closing; skeletonization; convolutional neural networks; hough transform; coordinate conversion;

INTRODUÇÃO

A agricultura brasileira evoluiu muito nas últimas décadas, graças às inovações e novas tecnologias, sendo responsável por 21% dos empregos formais no país, segundo Lamas (2023). Em 2022, a agricultura foi responsável por 24,8% do PIB e 47,6% das exportações, configurando-se como uma atividade de grande importância para o país. Para continuar com seu desenvolvimento, é necessário aperfeiçoar o uso de recursos, melhorando a produtividade e reduzindo desperdícios.

Nesse contexto, o uso de técnicas de georreferenciamento para o manejo permitem a realização de um controle mais preciso sobre o uso de recursos como combustível, defensivos agrícolas e fertilizantes. O georreferenciamento possibilita mapear detalhadamente as áreas cultivadas, identificando áreas de interesse que precisam de mais atenção e otimização no uso de insumos. Assim, obter os carregadores dentro de um campo de cultivo passa a ser uma tarefa necessária para o processo de automatização da agricultura.

A principal forma de obter esses dados atualmente é através de serviços de georreferenciamento utilizando drones ou imagens de satélite. Esses métodos fornecem dados precisos e atualizados, mas muitas vezes necessita de trabalho humano para a demarcação das posições dos carregadores, o que pode ser um processo oneroso e demorado.

O presente trabalho propõe-se utilizar uma metodologia semelhante à utilizada por K N et al. (2016), diferindo no uso de uma rede neural convolucional (CNN) e na identificação das posições geográficas das linhas de carregadores, visando acelerar o processo de demarcação e aumentar a precisão dos dados coletados, proporcionando um manejo mais eficiente e sustentável dos defensivos agrícolas.

MATERIAL E MÉTODOS

Materiais:

- Linguagem de programação de computadores Python versão 3.11;
- TensorFlow versão 2.13, biblioteca de código aberto para aprendizado de máquina e computação numérica;
- GeoJSON versão 3.1.0, biblioteca de codificação de estruturas de dados geográficos em python;
- PILLOW versão 10.4.0, biblioteca para manipulação e processamento de imagens;
- OpenCV versão 4.9.0.80, biblioteca de código aberto de software de visão computacional e aprendizado de máquina;
- Scikit-Image versão 0.23.2, conjunto de algoritmos para processamento de imagens;
- Albumentations versão 1.3.1, é uma biblioteca utilizada para aumento de imagem em tarefas de aprendizado profundo e visão computacional para melhorar a qualidade de modelos treinados;
- Software QGIS versão 3.30.2, aplicação de Sistema de Informações Geográficas gratuito e de código aberto que oferece suporte a visualização, edição e análise de dados geoespaciais;
- Computador com Placa de Vídeo Nvidia T4, 16GB de memória RAM e processador Intel Xeon CPU 2.20GHz.

Métodos:

A metodologia de trabalho para o desenvolvimento da pesquisa utilizou a sequência apresentada na Figura 1.

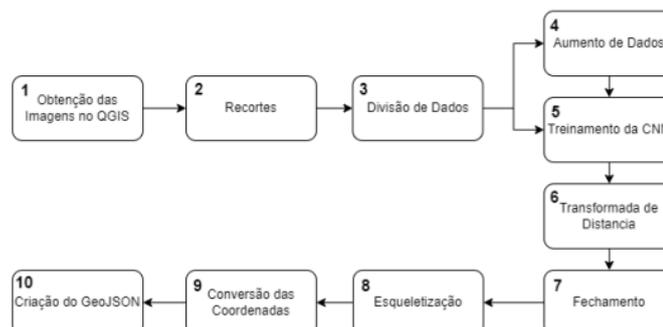


FIGURA 1. Sequencia da metodologia utilizada para o desenvolvimento deste trabalho.

1 - Obtenção das Imagens no QGIS: As imagens foram processadas através do QGIS para a demarcação do talhão, utilizando uma *bounding box* - um retângulo em que a área de interesse está inclusa - juntamente com as coordenadas do ponto superior esquerdo e inferior direito, obtendo assim a imagem contida na Figura 2.



FIGURA 2. Área de interesse destacada em vermelho, Imagem (A), e área exportada, Imagem (B) os pontos superior esquerdo e inferior direito do retângulo são exportados para utilização posterior.

Foi desenvolvido um algoritmo com o *Python*, que utiliza o interpretador disponibilizado pelo Qgis para a exportação das imagens e os pontos superiores e inferiores da *bounding box*, utilizados para calcular a diferença de latitude e longitude conforme as Equações 1 e 2.

$$\Delta lat = lat_{max} - lat_{min} \quad (1)$$

$$\Delta lon = lon_{max} - lon_{min} \quad (2)$$

2 - Recortes: As imagens de satélite possuem resolução e tamanho de arquivo variável de acordo com a *bounding box*, com dimensões próximas dos 30.000 pixels, adotou-se um método de fatiamento, reduzindo as imagens a dimensões máximas de 4096x4096 pixels conforme a Figura 3(A).

3 – Divisão dos Dados: Foram criados três conjuntos de imagens a partir das 2.376 imagens disponíveis. O conjunto de treinamento conta com 70% das imagens, o conjunto de validação com 20% e o conjunto de teste com os 10% restantes. O conjunto de validação é responsável por avaliar o desempenho do modelo a cada etapa do treinamento. Já as imagens do conjunto de teste não foram apresentadas à rede neural durante o processo de aprendizagem e são utilizadas exclusivamente para a realização de testes e verificação de métricas.

Os recortes foram preparados para treinar uma rede neural convolucional (*Convolutional Neural Network* - CNN) criada utilizando o *framework TensorFlow*, que possui diversas funções para lidar com fluxos de aprendizado de máquina. Para cada recorte da imagem original, há uma versão com as linhas de carregadores demarcadas. As demarcações das linhas foram realizadas manualmente utilizando o QGIS. O processo de recorte utilizado é o mesmo para as imagens com e sem demarcação. A Figura 3 apresenta um par de entrada/saída utilizados no treinamento da rede neural.

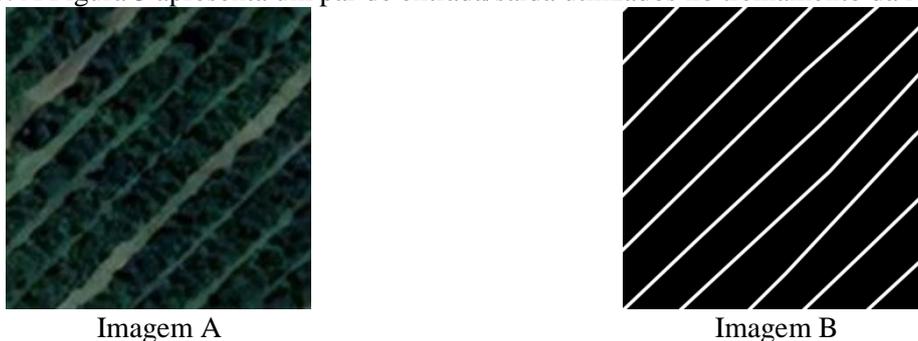


FIGURA 3. Exemplo de recorte, utilizado como entrada, Imagem A, e saída, Imagem B, esperada após o processamento pela Rede Neural.

4 – Aumento de dados: Para o conjunto de treinamento, foram realizadas operações de aumento de dados (*data augmentation*). Essas operações incluíram recortes, adições de ruídos, ajustes de brilho e contraste e rotações através da biblioteca *Albumentations*. As imagens resultantes das operações de aumento de dados possuem resolução de 2048x2048 pixels. Essas operações visaram à preparação das

imagens para o processamento posterior pela CNN garantindo que o modelo seja capaz de lidar com as variações e os ruídos presentes nos dados reais.

5 – Treinamento da CNN: O *TensorFlow* permite a construção de diferentes tipos de redes neurais, como as CNNs, utilizando modelos já estabelecidos como a *EfficientNetB0*, descrito por TAN; LE (2019), e a *U-Net*, documentada por RONNEBERGER; FISCHER; BROX (2015), além de permitir a construção de modelos próprios, neste trabalho foram utilizadas as duas arquiteturas de rede em conjunto para o desenvolvimento do projeto. O tamanho de entrada escolhido para a rede foi de 512x512 pixels, utilizando conexões de salto que permitem extrair diferentes características enquanto mantém a resolução da imagem.

Para a realização do treinamento, o objetivo foi minimizar a perda por *IoU* (*Intersection over Union*), também conhecido como índice de Jaccard. O *IoU* é uma métrica que visa medir a similaridade entre grupos de amostras. Como a rede neural deve segmentar as imagens de entrada, é desejável que sua saída seja a mais próxima possível da resposta esperada. O índice de Jaccard mede a similaridade entre a saída da rede e a resposta esperada, conforme a Equação 3.

$$IoU = \frac{T_p}{(T_p + F_p + F_n)} \quad (3)$$

em que,

T_p (*True Positives*) - pixels previstos pela rede como pertencentes à região que de fato o são;

F_p (*False Positives*) - pixels previstos pela rede como pertencentes à região, mas que não pertencem à região;

F_n (*False Negatives*) - pixels previstos pela rede como não pertencentes à região, mas que são pertencentes à região.

A função de perda utilizando o *IoU* é descrita na Equação 4.

$$Loss_{IoU} = 1 - IoU \quad (4)$$

Após o treinamento da CNN, o processamento subsequente das imagens pela rede resultou em representações visuais detalhadas e aprimoradas das linhas que representam os carregadores. Que podem ser observados na Figura 4(A), onde a imagem gerada demonstra as capacidades da rede em capturar e realçar as características desejadas, através do modelo em interpretar e transformar os dados visuais de entrada.

6 – Transformada de Distancia: A saída da rede neural foi submetida à Transformada de Distância (TD), que é uma operação utilizada para calcular a distância entre cada pixel e o pixel não nulo mais próximo. Essa técnica é essencial para identificar e quantificar a proximidade de diferentes regiões segmentadas na imagem, permitindo uma análise mais detalhada e precisa das áreas de interesse. A TD foi utilizada na tentativa de remover regiões distantes e desconexas do restante da imagem como demonstrado na Figura 4(B).

7 – Fechamento: A operação de fechamento é uma técnica de processamento de imagem utilizada para preencher pequenos buracos dentro de objetos na imagem. Esta operação é capaz de remover pontos isolados que estão fora do contexto da imagem principal, melhorando a continuidade e a integridade das regiões segmentadas.

Após a aplicação das etapas 6 e 7, foram obtidas imagens, conforme ilustrado na Figura 4(C). Este processo garantiu a clareza e a precisão dos detalhes na imagem, melhorando a qualidade e a nitidez do resultado final.

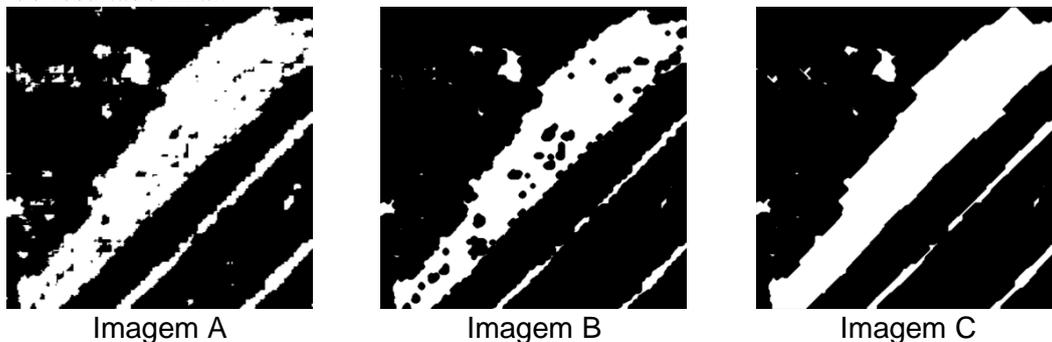


FIGURA 4. Imagens geradas pelo processamento da CNN.

8 – Esqueletização: A operação de esqueletização, disponível na biblioteca scikit-image, criou uma representação mais simplificada e linear da imagem, Figura 5, aproximando mais da linha real, “carreador”.

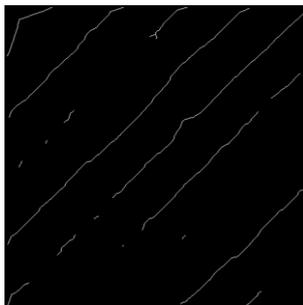


FIGURA 5. Exemplo de saída após o processo de esqueletização.

9 – Conversão das Coordenadas: Após a obtenção do esqueleto da imagem, foi aplicada a Transformada de Hough (TH), disponível na biblioteca OpenCV, para identificar os pontos inicial e final de cada linha na imagem. Com os pontos de início e fim de cada linha, juntamente com as posições georreferenciadas do ponto superior esquerdo, que corresponde a longitude e latitude mínimas, e inferior direito, representando a longitude e latitude máximas, correspondentes as posições dos pixels convertidas para coordenadas geográficas, utilizando as Equações 3 e 4:

$$x_{geo} = lon_{min} + \Delta lon \cdot \frac{x_{pixel}}{imglargura} \quad (3)$$

$$y_{geo} = lat_{min} + \Delta lat \cdot \left(1 - \frac{y_{pixel}}{imgaltura}\right) \quad (4)$$

As equações apresentadas acima fornecem uma conversão das coordenadas de pixel para coordenadas espaciais, desde que a área em questão possua variação inferior a 1°.

10 – Criação do GeoJSON: Após a conversão das coordenadas foi criado um arquivo GeoJSON, para ser importado por qualquer software de GIS. A estrutura do arquivo descrito na Figura 6 apresenta um exemplo de uma feição do tipo ponto com coordenadas específicas e um nome descritivo para a localização; o tipo do objeto, a geometria e a identificação da área de interesse.

```
{
  "type": "Feature",
  "geometry": {
    "type": "Point",
    "coordinates": [125.6, 10.1]
  },
  "properties": {
    "name": "Dinagat Islands"
  }
}
```

FIGURA 6. Exemplo de um arquivo GeoJSON descrevendo um ponto.

RESULTADOS E DISCUSSÃO

Como resultado do processamento, gerou-se um arquivo GeoJSON com as coordenadas espaciais das linhas de carreadores.

Em relação aos estudos de K N et al. (2016), este trabalho apresenta características semelhantes ao objetivo de identificar os carreadores conforme a metodologia aplicada. Ainda este estudo compartilha das mesmas dificuldades relatadas por Soares (2017) ao lidar com a TH, uma técnica amplamente utilizada para a detecção de linhas em imagens.

No entanto, o presente trabalho diferencia-se ao focar na obtenção da posição geográfica exata das linhas de carreadores, um objetivo que requer precisão adicional na análise e no processamento dos dados. Este enfoque geoespacial agrega um valor significativo, pois permite a representação das linhas em um contexto geográfico, facilitando sua visualização e utilização em sistemas de informação geográfica (GIS).

A Figura 7 foi o resultado, a partir do arquivo GeoJSON importado para o QGIS, proporcionando uma visualização clara e detalhada das linhas geográficas detectadas.

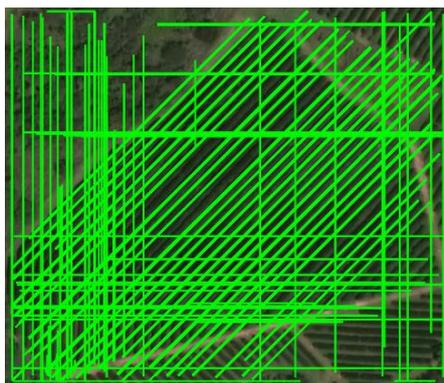


FIGURA 7. Visualização do arquivo GeoJSON gerado.

CONCLUSÕES

O presente trabalho conseguiu seu objetivo de forma parcial na obtenção dos dados georreferenciados. No entanto, a utilização da TH para a identificação das linhas apresentou falhas no processamento relacionado às curvas de nível, que resultaram em erros na sua detecção, que, nesse estudo, demonstrou-se limitada ao lidar com essas variações topográficas. Além disso, o algoritmo identificou múltiplas linhas onde deveria haver apenas uma, comprometendo a clareza e precisão dos resultados.

Conclui-se que, apesar de o trabalho ter atingido seus objetivos principais em termos de georreferenciamento, as dificuldades encontradas com a TH evidenciam a necessidade de aprimoramentos metodológicos, incluindo a integração de algoritmos complementares ou ajustes no pré-processamento dos dados para mitigar os problemas observados, especialmente em relação à detecção correta das linhas e à minimização de sobreposições indesejadas.

CONTRIBUIÇÕES DOS AUTORES

A.L.S.C. e H.G.S. contribuíram com a metodologia e testes. A.L.S.C. também foi responsável pela análise dos dados, criação da rede neural e algoritmos de PDI para reconhecimento dos padrões, além da redação do trabalho. H.G.S. participou do desenvolvimento dos algoritmos de PDI e análise dos dados. F.S.S. atuou na redação do trabalho e forneceu orientações para o desenvolvimento da metodologia.

Todos os autores contribuíram com a revisão do trabalho e aprovaram a versão submetida.

AGRADECIMENTOS

Agradecemos ao Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP) pelo suporte institucional e ao orientador pelo auxílio para o desenvolvimento deste trabalho.

REFERÊNCIAS

K N, R. et al. **Detection of Rows in Agricultural Crop Images Acquired by Remote Sensing from a UAV**. International Journal of Image, Graphics and Signal Processing, v. 8, n. 11, p. 25–31, 2016.

LAMAS, F. M. **A Evolução da Agricultura do Brasil**. Disponível em: <<https://www.embrapa.br/busca-de-noticias/-/noticia/81665485/artigo---a-evolucao-da-agricultura-do-brasil>>. Acesso em: jun. 2024.

RONNEBERGER, O.; FISCHER, P.; BROX, T. **U-Net: Convolutional networks for biomedical image segmentation**. arXiv (Cornell University), 2015.

SOARES, G. A. **Identificação de Linhas de Plantio por Meio da Transformada de Hough**. Faculdade de Computação Bacharelado em Sistemas de Informação: Universidade Federal de Uberlândia - UFU, 2017. Disponível em: <<https://repositorio.ufu.br/bitstream/123456789/19502/1/IdentificacaoLinhasPlantio.pdf>>. Acesso em: jun. 2024.

TAN, M.; LE, Q. V. **EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks**. 2019. Disponível em: <<http://arxiv.org/pdf/1905.11946>>. Acesso em: jun. 2024.