

Aplicando Redes Neurais Convolucionais em Imagens para Reconhecimento Automatizado de Abelhas Melíferas (*Apis mellifera* L.)

Caio Martim Barros¹, Emmanuel Diego Gonçalves de Freitas⁴,
Antonio Rafael Braga², Isac Gabriel Abrahão Bomfim³, Danielo G. Gomes¹

¹GREat, Departamento de Engenharia de Teleinformática - Centro de Tecnologia,
Universidade Federal do Ceara (UFC) – Fortaleza-CE

²Sistemas de Informação, Campus Quixadá,
Universidade Federal do Ceará, Quixadá-CE

³Departamento de Zootecnia, Setor de Abelhas,
Instituto Federal do Sertão Pernambucano (IF Sertão PE) - Santa Maria da Boa Vista-PE

⁴Departamento de Ensino, Área de Informática,
Instituto Federal do Ceará (IFCE), Iguatu-CE

caiobarros@alu.ufc.br, diego@ifce.edu.br

[rafaelbraga,danielo]@ufc.br, isac.bomfim@ifsertao-pe.edu.br

Abstract. *The recognition of bee species and other pollinators can directly contribute to the conservation of ecosystems. The objective of this work is to develop a model of digital image processing, for automatic recognition of honey bees (*Apis mellifera* L.) among the other bee species and other insects, through the use of convolutional neural networks. For training the proposed model, a data set consisting of 2.300 images was used, separated between the species of honey bee and other insects (including other bees in this class), which were called "Non Apis". The proposed neural network model returned answers with 94 % accuracy, culminating in a prediction model with a high precision index capable of recognize images of the same species of bee (*A. mellifera*) and differentiating it from other bees and other insects species.*

Resumo. *O reconhecimento de espécies de abelhas e demais polinizadores pode contribuir diretamente para a conservação de ecossistemas. O objetivo desse trabalho é desenvolver um modelo de processamento digital de imagens, para reconhecimento automático de abelhas melíferas (*Apis mellifera* L.) entre as demais espécies de abelhas e outros insetos, por meio do uso de redes neurais convolucionais. Para treinamento do modelo proposto foi utilizado um conjunto de dados composto por 2.300 imagens, separadas entre a espécie de abelha melífera e demais insetos (incluindo outras abelhas nessa classe), as quais foram denominadas de "Não Apis". O modelo de rede neural proposto respondeu com 94% de acurácia, culminando em um modelo de predição com um alto índice de precisão capaz de reconhecer imagens de abelhas melíferas e diferenciá-las das demais abelhas e outros insetos.*

1. Introdução

Para o ano de 2050, a previsão é que a população mundial alcance um número de nove bilhões de pessoas, o que demandará um incremento de até 60% na produção mundial de alimentos. Infelizmente, a expansão das fronteiras agrícolas ainda é uma das estratégias mais utilizadas para se elevar essa produção. Entretanto, essa é uma prática ambientalmente destrutiva, que leva ao amplo desmatamento das matas nativas, e, conseqüentemente, a irreparáveis perdas de biodiversidade [Potts et al. 2016]. Diante disso, a alternativa mais sensata, e a única sustentavelmente aceita, é aumentando a produtividade das áreas cultivadas. Para tal, seria necessário lançar mão de pesquisas, de caráter interdisciplinar, nos mais diversos âmbitos e níveis de tecnologia que foquem no melhor aproveitamento da terra. Assim, utilizando as mesmas áreas atualmente cultivadas, elevar-se-ia a produção de alimentos sem a necessidade de converter mais áreas naturais em áreas cultivadas [Freitas and Bomfim 2017].

Dentre os meios de se elevar a produtividade de plantios sem desmatar mais áreas, existe a introdução e o manejo de agentes polinizadores nas áreas de plantio [Milfont et al. 2013, Bomfim et al. 2016]. Outra forma é por meio da utilização de tecnologias para monitoramento e identificação precoce de pragas nos plantios [Gao et al. 2020, Iost Filho et al. 2020], como o Veículo Aéreo Não tripulado (VANT ou drone) e a visão computacional, que por sua vez possui a finalidade de detectar, segmentar, reconhecer e/ou identificar detalhes de regiões de interesse em uma determinada imagem, de modo automático ou semiautomático [Gonzalez and Woods 2011, Tian et al. 2020].

O uso de polinizadores com tecnologias da agricultura de precisão pode potencializar incrementos na produtividade das áreas cultivadas, minimizando assim a necessidade por mais desmatamentos. Além disso, a escolha da abelha melífera (*Apis mellifera* L.) como objeto de estudo se deve ao fato dela ser, dentre todas outras 20.000 espécies de abelhas catalogadas no mundo [Roskov et al. 2013], a mais utilizada na polinização da grande maioria das culturas agrícolas. Essa espécie é a preferida por, além da sua notável eficiência como polinizadora de uma gama de culturas agrícolas, ter seu criatório racional desenvolvido dentro de colmeias de fácil intervenção e transporte, bem como sua biologia e manejo amplamente estudados e difundidos [Moraes et al. 2012]. Entretanto, na visão de pessoas não treinadas, essa espécie de abelha pode ser confundida com outras abelhas, vespas e moscas ou mesmo insetos-praga, o que poderia levar a falha no acompanhamento dos seus serviços de polinização dentro de plantios [Mason et al. 2018].

Nesse contexto, as tecnologias desenvolvidas para monitorar a atividade das campeiras de colônias de abelhas melíferas introduzidas em plantios podem vir a auxiliar nas tomadas de decisão, visando melhorar a densidade de polinizadores e seu direcionamento nas flores. Uma alternativa, para o alcance de tais objetivos, é o desenvolvimento de sistemas de visão computacional, baseados em Processamento Digital de Imagens (PDI) e métodos de inteligência artificial, capazes de reconhecer automaticamente a abelha melífera dentre os demais insetos benéficos ou pragas que podem estar presentes no plantio.

O objetivo deste trabalho, portanto, é gerar um modelo computacional de visão computacional, para reconhecimento automático de abelhas melíferas (*Apis mellifera* L.) entre as demais espécies de abelhas e outros insetos, por meio do uso de redes neurais convolucionais (CNN - *Convolutional Neural Network*), propondo analisar o desempenho do modelo construído com base em uma outra arquitetura conhecida no ramo de imagens,

que é a rede Neural MobileNet, comparando os parâmetros de acurácia (*accuracy*), que esta será representada em porcentagem pois os valores estão entre 0 e 1, e perda (*loss*), a qual será representada por valores decimais.

2. Trabalhos Relacionados

A literatura apresenta trabalhos relativos à segmentação de imagens de asas de abelhas com técnicas de limiarização, dilatação, erosão, afinamento (ou esqueletização), *hit-and-miss* e, para redução de ruído e técnicas de filtragem para a preservação de borda [Fagundes et al. 2020]. Uma técnica que tem sido bastante utilizada para processamento e análise de imagens digitais são as redes neurais convolucionais (*Convolutional Neural Networks - CNN*), em que encontramos contribuições para classificação de espécies de abelhas e de borboletas através das imagens de suas asas [Liu et al. 2020] [Buschbacher et al. 2020].

Insetos em milho, soja e trigo tem sido detectados e classificados com o uso de algoritmos de aprendizagem de máquina [Kasinathan et al. 2020], como também vespas (*Anisopteromalus calandrae*) e (*Heterospilus prosopidis*) e besouros (*Callosobruchus chinensis*) [Tuda and Luna-Maldonado 2020].

Para avaliar a saúde de abelhas, podemos destacar o uso de técnicas de visão computacional para classificar abelhas como 'saudáveis' e 'infectadas' a partir da presença/ausência do ácaro ectoparasita *Varroa destructor* [Schurischuster and Kampel 2020]. As imagens do referido artigo de Schurischuster e Kampel (2020) contêm todas as partes do corpo das abelhas, assim como as imagens utilizadas para os experimentos deste trabalho.

3. Material e Métodos

3.1. Conjunto de dados

Para geração do modelo proposto, foi utilizado um conjunto de dados composto por 2.300 imagens de tamanhos variados (largura entre 25 e 1.500 pixels, altura entre 20 e 1.000 pixels), das quais 1.150 contam com a presença de abelhas melíferas (*Apis mellifera*), enquanto as outras 1.150 estabelecem um grupo de insetos diversos, aqui denominados como "Não Apis". A Figura 3.1 apresenta amostras de imagens do conjunto de dados.

A construção do conjunto de dados foi feita a partir de uma revisão de bases disponíveis para o objetivo proposto neste trabalho. Duas bases foram encontradas publicamente: (1) o conjunto de dados *BeeImage* [Yang 2019], com mais de 5.100 imagens de abelhas rotuladas com localização, data, hora, subespécie, condição de saúde, casta e pólen; (2) O conjunto de dados *Bee or wasp?* [Rey 2020], criada em 2020 para aplicações com objetivo de distinguir abelhas de vespas. Esse conjunto possui um total de 7.942 imagens, onde 3.183 são rotuladas como abelhas, 4.943 como vespas, 2.453 como outros insetos e 845 rotuladas como outros objetos.

A partir da base de [Yang 2019] foram selecionadas 1.110 imagens de abelhas melíferas (*A. mellifera* L.), enquanto da base de [Rey 2020] foram 1.120 imagens de insetos diversos (outras espécies de abelhas, vespas, moscas e outros insetos), referidos aqui como "Não Apis". Em ambas as bases, o processo de seleção de imagens foi realizado por um apidólogo (estudioso das abelhas). Ademais, para se alcançar um número de 1.150

imagens para cada uma das duas classes (*Apis mellifera* e "Não Apis"), foram acrescentadas mais 40 e 30 imagens, respectivamente, sendo todas elas coletadas por meio da Internet e também validadas pelo apidólogo.

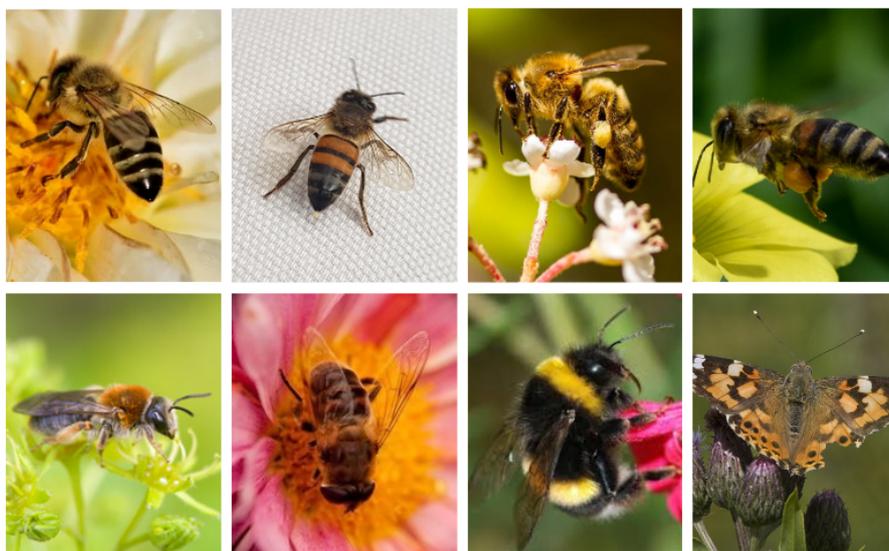


Figura 1. Amostras de imagens utilizadas: as imagens na parte superior são de *Apis mellifera*, enquanto as imagens na parte inferior são de "Não Apis". Fonte: [Yang 2019] e [Rey 2020].

3.2. Redes Neurais Convolucionais

Uma rede neural convolucional é definida por diversos hiper-parâmetros que descrevem o formato geral do modelo de aprendizagem de máquina, a citar: o número de camadas, número de filtros por camada e número total de parâmetros. Com base nisso, [Freitas et al. 2019] ressaltam que o processo de desenvolvimento de uma rede neural como solução de novos problemas é um desafio frequente, sobretudo na busca da arquitetura de melhor abstração e generalização, pois a parametrização da arquitetura é um processo exploratório e não definitivo. No caso do PDI e sua classificação, a rede convolucional é uma das arquiteturas de redes neurais mais usadas para a identificação de espécies, superado de forma mais satisfatória diversos métodos tradicionais, tais como o *Random Forest* (RF) e o *Support Vector Machine* (SVM) [Monteiro et al. 2019].

Na Figura 3.2 é ilustrado um diagrama com as principais camadas utilizadas em uma CNN. É possível sequenciar quantas camadas convolucionais forem necessárias, com ou sem camadas de *pooling* (para a redução de escala), antes que seja feita a operação de *flatten*, cujo objetivo é sequenciar todos os dados processados em uma lista unidimensional para a camada densa.

Para o objetivo proposto, foi elaborado um modelo composto por 3 camadas convolucionais, onde cada uma delas é seguida por outra camada de *pooling*, conforme a representação ilustrada pela Figura 3.2.

O modelo desenvolvido foi baseado no padrão convencional de uma CNN, conforme apresentado na Figura 3.2, contudo, alguns parâmetros foram variados na tentativa

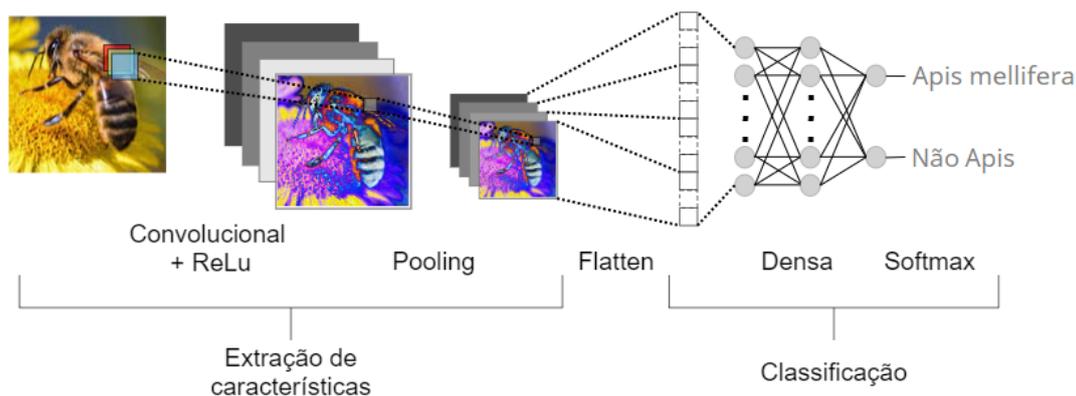


Figura 2. Exemplo de uma rede neural CNN para a detecção de imagens de abelhas.



Figura 3. Arquitetura do modelo de CNN desenvolvido.

de obter um modelo que melhor se adapte ao problema em questão. As variações de modelos consideraram camadas convolucionais que poderiam possuir 32, 64 e 128 filtros, que por sua vez poderiam ter dimensões de área iguais a 3x3, 5x5 e 9x9. As condições desses parâmetros foram testadas em diversas combinações possíveis, com o propósito de encontrar qual delas teria melhor resultado dentro do contexto estudado. Na Figura 3.2, também são mostrados os parâmetros selecionados para cada camada do modelo proposto, seguindo a supracitada metodologia de escolha.

O modelo baseado em CNN foi desenvolvido e treinado com a API Keras, para a biblioteca TensorFlow da Google, utilizando a linguagem Python. Nesse modelo, a camada de entrada foi configurada para receber imagens de 3 canais com 32x32 pixels. Para execução dos testes propostos na metodologia adotada, foi utilizado um computador com processador AMD Ryzen 3 3200G 3.60GHz, 16 GB de RAM e Placa de Vídeo RX 570 8GB, gerenciado pelo Sistema Operacional Windows 10 Professional.

3.3. Treinamento da CNN

A base teve 70% das imagens selecionadas, aleatoriamente, para treino e 30% para testes. Também foi criado um terceiro conjunto de dados, composto por 20% de imagens selecionadas aleatoriamente do conjunto de treino, com o propósito de realizar a validação.

Como o processo de treinamento é, basicamente, um tarefa de otimização de uma função custo, o aprendizado do modelo está condicionado ao desempenho do método de otimização. Como a finalidade de analisar qual o melhor otimizador dentro do contexto criado pelo objetivo deste trabalho, foram testadas quatro otimizadores: adam, adagrad, adadelta e RMSprop. Na Tabela 3.3 são expostos os resultados obtidos com cada otimiz-

dor testado, utilizando uma arquitetura igual para todos os testes. Observando os dados, percebe-se que o otimizador adam tem o melhor desempenho para esse tipo de análise, fato adotado como critério para sua escolha como otimizador no treinamento do modelo que foi desenvolvido.

Tabela 1. Resultados dos otimizadores usando o modelo padrão

Otimizadores	Acurácia	Perda
Adam	0.9405	0.1490
RMSprop	0.9246	0.1944
Adagrad	0.6797	0.5771
Adadelta	0.7159	0.6832

4. Resultados

4.1. Resultados com o modelo proposto

O modelo proposto, ilustrado na Figura 3.2, foi testado com o objetivo de se observar as métricas de avaliação acurácia, *f1-score*, *recall* e precisão (Tabela 4.1). Como forma de observar o desempenho do treinamento, para avaliação dos parâmetros utilizados nesse processo, também foram coletadas as respostas da função custo (*loss* ou perda), durante as épocas de treino. Na Figura 4.1 é possível avaliar a evolução da perda, durante o treinamento que trouxe 0.94 (ou 94%) de acurácia, utilizando os parâmetros selecionados a partir da metodologia proposta.

Tabela 2. Resultados do teste com o melhor modelo.

	precision	recall	f1-score	support
Apis	0.97	0.91	0.94	345
NaoApis	0.92	0.97	0.94	345
accuracy			0.94	690
macro avg	0.94	0.94	0.94	690
weighted avg	0.94	0.94	0.94	690
Loss function: 0.14 accuracy: 0.94				

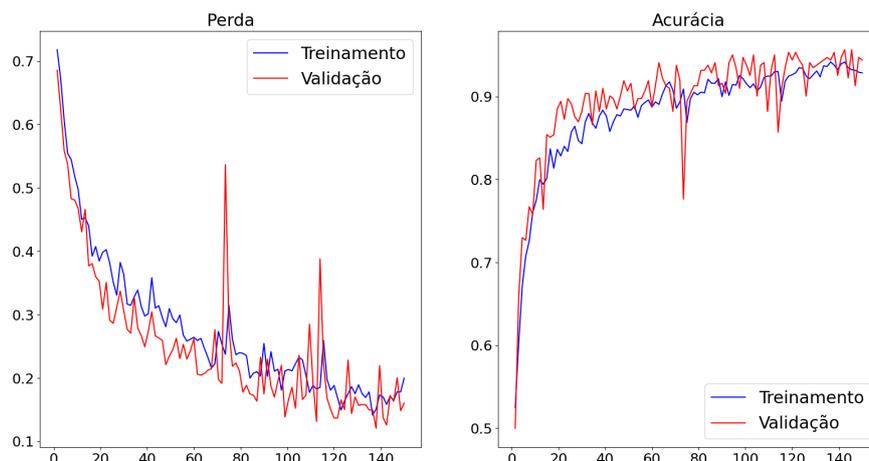


Figura 4. Acurácia e perda do modelo proposto.

4.2. Resultados com a CNN MobileNet

Para efeito de comparação do modelo criado, foi treinada também uma arquitetura amplamente usada na área de PDI, a MobileNet ¹. Nessa etapa, adaptou-se o modelo criado para a CNN de acordo com a Figura 3.2, seguindo os parâmetros necessários para a configuração do modelo via rede neural MobileNet, onde é alocado um *dropout* de *rate* 0,25, usando o otimizador Adam para compilar, os quais restantes parâmetros são padrões para essa rede. Por fim, gerou-se o gráfico de resultados vigentes desse teste, a qual pode ser analisada pela Figura 4.2, em que se encontram valores 0.29 de perda e 0.89 de acurácia.

Tabela 3. Resultados do teste utilizando MobileNet.

	precision	recall	f1-score	support
Apis	0.88	0.91	0.90	863
NaoApis	0.91	0.88	0.89	862
accuracy			0.90	1725
macro avg	0.90	0.90	0.90	1725
weighted avg	0.90	0.90	0.90	1725
Loss function: 0.29 accuracy: 0.89				

5. Discussão

Embora a diferença entre o modelo obtido neste trabalho e a arquitetura MobileNet revele bons resultados, no tocante ao objetivo geral da pesquisa, uma nova revisão foi construída a partir da observação de resultados em artigos relevantes para o tema e citações com valores de resposta para aplicações com *CNN*, em tópicos relacionados à abordagens de PDI de abelhas. Com essa abordagem, foi observada alta frequência de taxas acima de 90% de

¹<https://keras.io/api/applications/mobilenet/>

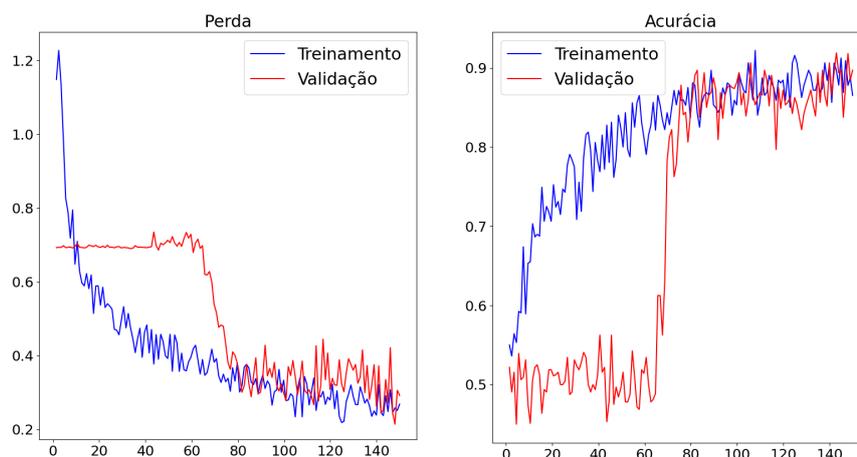


Figura 5. Acurácia e perda do modelo MobileNet.

acurácia, como pode ser analisado em [Sledevič 2018], no qual é realizado um estudo sobre a abelha *Bombus* (uma espécie de abelha diferente da *Apis mellifera*), utilizando uma rede neural convolucional que trouxe respostas com taxas de acurácia entre 92% a 94%. Em [Schurischuster and Kampel 2020], uma nova arquitetura, dita DeepLabv3, com base na ResNet-101, que é uma rede neural convolucional, é desenvolvida para a construção de um modelo de detecção de varroas, cujos os testes trouxeram respostas com taxas de acurácia superiores a 90%. Também observou-se que os trabalhos relacionados mencionados não costumam ressaltar o valor do parâmetro *loss*.

Sendo assim, a partir do estudo apresentado dos métodos propostos, foi possível obter um modelo de predição com 94% de acurácia com a CNN. Este resultado é observado como satisfatório, principalmente quando comparado com a CNN MobileNet que, quando testada com a mesma base, conseguiu uma acurácia de 89,5%. A partir de uma análise comparativa, feita sobre os resultados do modelo proposto e dos resultados em trabalhos relacionados, é possível perceber que as respostas obtidas neste trabalho pertencem à faixa dada nas tabelas referenciais oriundas de trabalhos relacionados. Todavia, a taxa de perda ao final do treinamento, que é 0.15 para o modelo proposto, ainda pode ser investigada para fins de melhoramento, como objetivo de conseguir percentagens menores, tais como 0.01 ou inferior.

6. Conclusão

Este trabalho inova trazendo um reconhecimento automático com alta acurácia de imagens de abelhas melíferas (*A. mellifera*), diferenciando-as, assim, das imagens dos demais insetos costumeiramente encontrados em plantios. Desta forma, entende-se que a abordagem proposta traz uma contribuição pioneira e relevante, visto que a parametrização e aplicação de CNN em distinção entre abelhas melíferas (*Apis mellifera*) e outros insetos, ainda é um tema pouco explorado, segundo a revisão feita para o desenvolvimento desta pesquisa.

Como trabalhos futuros, os autores pretendem utilizar *transfer learning* em conjunto com a CNN e também outras arquiteturas de redes neurais, bem como a utilização de Redes Neurais com Funções de Ativação de Base Radial, com o objetivo de melhor

definir a arquitetura de rede neural para esse fim de classificação de espécies de abelhas. Pretende-se também ajustar todas as imagens disponíveis a fim de retirar o plano de fundo contidas nelas, da forma que seja possível a realização das predições das imagens somente com foco nas características das abelhas. Por fim, após a melhoria da arquitetura de detecção e diferenciação entre as abelhas melíferas dos demais insetos, deseja-se avançar na tecnologia de sensoriamento remoto da atividade das campeiras de colônias de abelhas melíferas introduzidas em plantios, visando auxiliar no manejo de polinização dessas áreas para incremento da produtividade agrícola e, assim, colaborar sustentavelmente com aumento da produção de alimentos.

Agradecimentos

O presente artigo foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Danielo G. Gomes agradece o suporte financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) através dos processos 432585/2016-8 e 310317/2019-3.

Referências

- Bomfim, I. G. A., Freitas, B. M., de Aragão, F. A. S., Walters, S. A., and Pessarakli, M. (2016). Pollination in cucurbit crops. *Handbook of Cucurbits: Growth, Cultural Practices, and Physiology. Florida (US): CRC Press. <https://doi.org/10.1201/b19233-16>*.
- Buschbacher, K., Ahrens, D., Espeland, M., and Steinhage, V. (2020). Image-based species identification of wild bees using convolutional neural networks. *Ecological Informatics*, 55:37 – 45.
- Fagundes, J. M. G., Rebelo, A. R., Digiampietri, L. A., and Bísvaro, H. H. (2020). Fully automatic segmentation of bee wing images. *Revista Brasileira de Computação Aplicada*, 12 (2):37 – 45.
- Freitas, B. M. and Bomfim, I. G. A. (2017). A necessidade de uma convivência harmônica da agricultura com os polinizadores. *Importância dos polinizadores na produção de alimentos e na segurança alimentar global. Brasília, DF: Centro de Gestão e Estudos Estratégicos*.
- Freitas, D., Camargo, S., Comin, H., Domingues, R., Gaspar, E., and Cardoso, F. (2019). Reconhecimento da ceratoconjuntivite infecciosa bovina utilizando imagens termográficas e redes neurais convolucionais. *Revista Brasileira de Computação Aplicada*, 11(3):133–145.
- Gao, D., Sun, Q., Hu, B., and Zhang, S. (2020). A framework for agricultural pest and disease monitoring based on internet-of-things and unmanned aerial vehicles. *Sensors*, 20(5):1487.
- Gonzalez, R. C. and Woods, R. E. (2011). *Digital Image Processing*. Pearson Prentice Hall, 3th edition.
- Iost Filho, F. H., Heldens, W. B., Kong, Z., and de Lange, E. S. (2020). Drones: Innovative technology for use in precision pest management. *Journal of economic entomology*, 113(1):1–25.

- Kasinathan, T., Singaraju, D., and Uyyala, S. R. (2020). Insect classification and detection in field crops using modern machine learning techniques. *Information Processing in Agriculture*.
- Liu, S., Shih, F. Y., Russell, G., Russell, K., and Phan, H. (2020). Classification of ecological data by deep learning. *International Journal of Pattern Recognition and Artificial Intelligence*, 34 (13):37 – 45.
- Mason, L., Kondratieff, B. C., and Seshadri, A. (2018). *Native Bee Watch: A Colorado Citizen Science Field Guide*. Colorado State University, College of Agricultural Sciences.
- Milfont, M. d. O., Rocha, E. E. M., Lima, A. O. N., and Freitas, B. M. (2013). Higher soybean production using honeybee and wild pollinators, a sustainable alternative to pesticides and autopollination. *Environmental chemistry letters*, 11(4):335–341.
- Monteiro, A. A. d. O. et al. (2019). Segmentação e identificação de cultivo e espécies invasivas em imagens aéreas utilizando redes neurais convolucionais.
- Morais, M. M., De Jong, D., Massage, D., and Gonçalves, L. S. (2012). Perspectivas e desafios para o uso das abelhas *apis mellifera* como polinizadores no brasil. *Polinizadores no Brasil. Edusp, São Paulo, Brasil*.
- Potts, S. G., Imperatriz-Fonseca, V., Ngo, H., Biesmeijer, J. C., Breeze, T., Dicks, L., Garibaldi, L., Settele, J., Vanbergen, A. J., Aizen, M. A., et al. (2016). Summary for policymakers of the assessment report of the intergovernmental science-policy platform on biodiversity and ecosystem services (ipbes) on pollinators, pollination and food production. Technical report.
- Rey, G. (2020). The bee or wasp? dataset: 19480 hand curated photos of bees, wasps and other insects. <https://www.kaggle.com/jerzydziewierz/bee-vs-wasp>. Accessed: 2021-04-12.
- Roskov, Y., Kunze, T., Paglinawan, L., Orrell, T., Nicolson, D., Culham, A., Bailly, N., Kirk, P., Bourgoin, T., Baillargeon, G., et al. (2013). Species 2000 & itis catalogue of life, 2013 annual checklist.
- Schurischuster, S. and Kampel, M. (2020). Image-based classification of honeybees. In *2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–6.
- Sledevič, T. (2018). The application of convolutional neural network for pollen bearing bee classification. In *2018 IEEE 6th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, pages 1–4.
- Tian, H., Wang, T., Liu, Y., Qiao, X., and Li, Y. (2020). Computer vision technology in agricultural automation —a review. *Information Processing in Agriculture*, 7(1):1–19.
- Tuda, M. and Luna-Maldonado, A. I. (2020). Image-based insect species and gender classification by trained supervised machine learning algorithms. *Ecological Informatics*, 60:101135.
- Yang, J. (2019). The beeimage dataset: Annotated honey bee images. <https://www.kaggle.com/jenny18/honey-bee-annotate-images>. Accessed: 2021-02-20.