

Detecção e Identificação de Pólen em Imagens de *Apis mellifera* por Meio de Redes Neurais Convolucionais

João P. S. Cardoso^{1(*)}, Manoel F. C. Neto¹, Roberto C. L. Oliveira¹, Marcos E. C. Oliveira², Daniel S. Pereira²

¹Instituto de Tecnologia – Universidade Federal do Pará (UFPA)

²Embrapa Amazônia Oriental - (CEPATU)

{(*)jpscaldoso, limao}@ufpa.br, manael.campos@itec.ufpa.br, {marcos-ene.oliveira, daniel.pereira}@embrapa.br

Resumo. *O estudo visa identificar abelhas Apis mellíferas carregando pólen através de imagens da entrada da colmeia, usando redes neurais convolucionais (CNN). Isso possibilita o monitoramento automatizado das colmeias e coleta de dados sobre o comportamento das abelhas em larga escala. A pesquisa aplicou a técnica de transferência de aprendizado na CNN AlexNet, alcançando uma impressionante acurácia de 99.3% e um F-Score de 99.29% na classificação das imagens entre aquelas com e sem pólen.*

1. Introdução.

As abelhas desempenham um papel crucial na polinização, vital tanto para a agricultura quanto para a vida humana, sendo objeto de estudo há quase um século [Lundie 1925]. No entanto, observar suas atividades de maneira tradicional, com observação humana e anotações manuais, é demorado e invasivo, podendo perturbar a colmeia e afetar seus fenômenos. A observação natural de colmeias de abelhas é valiosa em várias aplicações, incluindo a prevenção de doenças, a detecção de materiais tóxicos como pólen adulterado e o diagnóstico da saúde da colmeia [Schurischuster 2016] [Frias 2016].

O uso de tecnologias avançadas, como visão computacional e aprendizado de máquina, facilita a observação e proporciona novos *insights* sobre as dinâmicas sociais das colmeias [Krizhevsky 2012]. Este trabalho propõe o reconhecimento de abelhas com pólen em imagens por meio de redes neurais convolucionais (CNN). Serão apresentadas as CNNs e a rede *Alexnet* nas próximas seções, seguidas pela metodologia, conjunto de dados, técnica de transferência de aprendizado, hiperparâmetros, resultados e conclusões.

2. Materiais e Métodos.

2.1 Redes Neurais Convolucionais.

As Redes Neurais Convolucionais (CNNs) são uma variação das redes de *Perceptrons* de Múltiplas Camadas projetada para processar dados visuais, tratando cada entrada como uma imagem. Uma característica fundamental das CNNs é a preservação da vizinhança entre *pixels* durante o processamento [Lecun 2010]. Elas são compostas por camadas especializadas, incluindo camadas convolucionais para extração de atributos, camadas de *pooling* para redução de dimensões e camadas

totalmente conectadas para propagar o sinal e produzir probabilidades de classificação em relação às classes previamente treinadas. Nos últimos anos, as redes neurais convolucionais (CNNs) se destacaram como a abordagem de ponta para classificação visual [Krizhevsky 2012].

2.2 AlexNet.

A *AlexNet*, desenvolvida por [Krizhevsky 2012], é uma CNN projetada para tarefas de classificação visual. Ela se destacou ao vencer o desafio de reconhecimento visual em grande escala da *ImageNet* (ILSVRC) e se tornou amplamente utilizada no campo do *deep learning* [Shin 2016]. A arquitetura da *AlexNet*, inclui cinco camadas convolucionais e três camadas totalmente conectadas. A última camada totalmente conectada classifica em 1000 classes, enquanto o restante da rede age como um extrator de *features*, gerando vetores de *features* de 4096 dimensões para cada imagem, incluindo ativações antes da camada de saída. A arquitetura de uma *AlexNet* possui: 5 camadas convolucionais, 3 de *max-pooling*, 2 normalizadas, 2 camadas totalmente conectadas e 1 camada *SoftMax* [Krizhevsky 2012].

2.3 Metodologia.

Nesta seção, descreveremos o conjunto de dados usado para treinar o sistema de classificação baseado na CNN *AlexNet*, que categoriza imagens em "Com Pólen" ou "Sem Pólen". Usamos a técnica de transferência de aprendizado, substituindo camadas de classificação, e aplicamos aumento de dados com translações nos eixos x e y. Foi utilizado o software *Matlab* e o *Deep learning toolbox*. Dividiu-se o trabalho em três etapas: criação do *dataset*, treinamento da rede e análise dos resultados para avaliar o desempenho.

2.4 Conjunto de dados.

Este conjunto de dados de imagens foi criado a partir de vídeos capturados na entrada de uma colônia de abelhas em junho de 2017 nas instalações de abelhas da Estação Experimental Agrícola de Gurabo da Universidade de Porto Rico. [Rodriguez et al 2018]. Contém imagens de alta resolução de abelhas portadoras e não portadoras de pólen. O conjunto de dados possui 700 imagens de tamanho 180 x 300 pixels, todas as imagens são de extensão do tipo JPG. Os dados foram divididos em dados de treinamento, validação e teste, correspondendo respectivamente, a 60%, 20% e 20% do total das imagens, esta escolha se deu como pode ser observado na Tabela 1.

Tabela 1: Divisão dos dados para o modelo.

Classe	Treino	Validação	Teste	Total
Com Pólen	212	70	72	354
Sem Pólen	207	70	69	346
Geral	419	140	141	700

2.5 Transferência de Aprendizado.

A transferência de aprendizado é uma técnica amplamente empregada no campo do *deep learning*, visando ajustar uma CNN pré-treinada para realizar a classificação em uma nova coleção de imagens. Esta abordagem oferece vantagens em termos de eficiência, uma vez que é mais rápida e menos exigente em termos de volume de imagens de treinamento em comparação com o treinamento do zero [Karpathy 2018]. Para realizar a transferência, basta substituir o bloco de camadas relacionadas à classificação, que inclui as três últimas camadas: totalmente conectada, *Softmax* e de classificação. Desta forma, foi configurado a primeira camada para imagens de 227 x 227 x 3. Em seguida foi substituído as três últimas camadas e transferido 22 camadas, totalizando 25. Os hiper parâmetros de configuração da rede utilizada foram determinados por meio de experimentação, como detalhado na Tabela 2.

Tabela 2 Configurações de hiper parâmetros.

Otimização	Descida gradiente estocástica com otimização de impulso (SGDM).
MiniBatchSize	64
MaxEpochs	20
initialLearnRate	0.001
ValidationData	augimdsValidation (aumento de dados de validação.)
ValidationFrequency	50

3. Resultados

A Figura 1 ilustra os testes de classificação realizados com a CNN, após a transferência de aprendizado e novo treinamento.

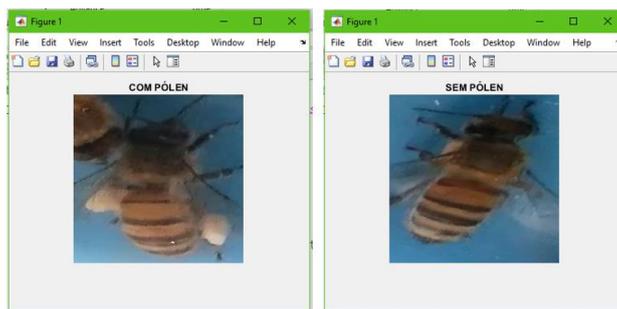


Figure 1 Classificação realizada com imagem de teste, utilizando a CNN com base na AlexNet. Fonte: Autores.

A acurácia ao longo das iterações, teve como tempo 8 min e 3 seg, atingindo o valor máximo de 99,29%. A Figura 2 ilustra a matriz de confusão obtida pós treinamento. A matriz indica uma acurácia geral de 99,3%. Os resultados obtidos através dos dados de teste, as métricas de avaliação utilizadas foram: sensibilidade (Sens), precisão (Prec) e *F-score*. Na Figura 2 observa-se uma maior sensibilidade na classe sem pólen e uma maior precisão na classe com pólen. Estas medidas indicam boa compreensibilidade dos dados de entrada, com quase totalidade na classificação correta das imagens. A rede obteve ótimos resultados gerais de sensibilidade, precisão e *F-score*, com esta configuração.

		Confusion Matrix		
		COM PÓLEN	SEM PÓLEN	Total
Output Class	COM PÓLEN	72 51.1%	0 0%	100% 0.0%
	SEM PÓLEN	1 0.7%	68 48.2%	98.6% 1.4%
		COM PÓLEN	SEM PÓLEN	Total
		98.6% 1.4%	100% 0.0%	99.3% 0.7%
		Target Class		

Figura 2 Matriz de confusão para os dados de teste. Fonte: Autores.

4. Conclusões

Neste trabalho foi apresentado um modelo para sistema de classificação de imagens utilizando CNN, com o objetivo de identificar pólen em imagens de abelhas coletadas de um *dataset*. Pode ser destacado em relação ao sistema, a utilização de rede pré-treinada como base, o que diminui o custo computacional de treinamento. Além disso o sistema foi capaz de realizar previsões corretas, mesmo utilizando uma base de dados considerada pequena, tendo acurácia de 99.3%.

Referências

- Frias, B. E. D., Barbosa, C. D., & Lourenço, A. P. "Pollen nutrition in honey bees (*apis mellifera*): impact on adult health." *Apidologie*, 47(1), 15–25, 2016.
- Ivan Rodriguez, Rémi Mégret, Edgar Acuña, José Agosto, Tugrul Giray. Recognition of pollen-bearing bees from Video using Convolutional Neural Network, (accepted at) IEEE Winter Conf. on Applications of Computer Vision, 2018, Lake Tahoe, NV.
- Karpathy, A. "Transfer learning and fine-tuning convolutional neural networks." <http://cs231n.github.io/transfer-learning/>. 2015. (Acessado em 03/07/2018).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. "ImageNet classification with deep convolutional neural networks." In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25*, pp. 1097–1105. Curran Associates, Inc., 2012.
- LeCun, Y., Kavukcuoglu, K., & Farabet, C. "Convolutional networks and applications in vision: Proceedings of 2010 IEEE International Symposium on Circuits and Systems, pp. 253–256." (2010).
- Lundie, A. E. "The flight activities of the honeybee." 1925.
- Schurischuster, S., Zambanini, M., Kampel, M., & Lamp, B. "Sensor study for monitoring varroa mites on honey bees (*apis mellifera*)." *Proceedings of the Visual Observation and Analysis of Vertebrate and Insect Behavior (VAIB) Workshop*, 2016.
- Shin, H. C., et al. "Redes neurais convolucionais profundas para detecção auxiliada por computador: arquiteturas da CNN, características do conjunto de dados e aprendizado de transferência." *Transações do IEEE em imagens médicas* 35.5 (2016), pp. 1285-1298.