XXXII Encontro Anual de Iniciação Científica
IX Encontro Anual de Iniciação Científica Jr.
EAIC 2023

PROCESSO DE INDEXAÇÃO DE IMAGENS DE DOENÇAS EM PLANTAS

Cleberton Kobay dos Santos (PROVIC), Emili Everz Golombiéski, Maria Salete Marcon Gomes Vaz, e-mail: salete@uepg.br.

Universidade Estadual de Ponta Grossa/Departamento de Informática

Ciência da Computação, Metodologia e Técnicas de Computação.

Palavras-Chave: *Banco de Dados, Classificação de doenças agrícolas, Culturas Agrícolas.*

Resumo

Metadados são utilizados para indexar e catalogar diversos tipos de informações. O objetivo desta pesquisa foi indexar imagens de doença em plantas, extraindo os metadados e analisando a estrutura da planta. Para tanto, foi definido um processo de indexação, incluindo a definição de um algoritmo e a cultura agrícola. Como resultados, houve a captação e a geração de metadados mais rápidas e eficientes.

Introdução

A indexação permite a busca de imagens semelhantes com base em características visuais (metadados).
A indexação de imagens desempenha um papel importante na análise de doenças em plantas, proporcionando aos agricultores uma ferramenta eficaz para detectar patógenos e melhorar a produção agrícola.
Esta pesquisa teve como objetivo a indexação de imagens de doenças de plantas, com foco nas ferrugens de milho, aplicando métodos de aprendizado de máquina.

Materiais e Métodos

Para realizar esta pesquisa, foram utilizados dois bancos de imagens, *Corn or Maize Leaf Disease Dataset* e o *Corn Leaf Infection Dataset,* ambos de folhas de milho, contendo doenças de ferrugens comuns e a doença *Turcicum*. Os bancos contém 3562 imagens. Para pré-processamento de imagens foi utilizada a Biblioteca Keras (Keras, 2023), uma biblioteca de aprendizado de máquina desenvolvida em Python. Também, foi utilizada uma rede neural convolucional (CNN) para classificar e indexar imagens, implementada com o Framework *TensorFlow* (TensorFlow, 2023). A arquitetura da rede consistia de três camadas convolucionais, A função de perda utilizada foi a *Categorical Crossentropy*, e a otimização foi feita pelo Algoritmo *Adam*.
O CBIR (*Content Based Image Retrieval*) foi utilizado para recuperação de imagens baseada em conteúdo.

Resultados e Discussão

Uma rede neural convolucional foi treinada para classificar e indexar as imagens de folhas de milho com alta precisão, implementada TensorFlow.
Inicialmente, o modelo apresentou um problema de *overfitting, com* memorização dos dados de teste. O modelo não foi capaz de identificar dados novos fora do padrão da base de testes. Esse problema foi solucionado com a melhoria da base de dados, com embaralhamento dos dados de treino.
Os resultados do treinamento da CNN alcançaram uma acurácia de 0.9984 e perda de 0.0079 no conjunto de treinamento e 1.2807 de perda e 0.7757 de precisão na validação, demonstrando que o modelo aprendeu as características (metadados) de doenças de plantas.
Para melhoria da eficácia da indexação, foi introduzido o Método CBIR (*Content Based Image Retrieval*) para recuperação de imagens baseada em conteúdo.
Nas primeiras execuções não apresentou uma acurácia adequada, a qual foi solucionada com o ajuste dos parâmetros e otimização do cálculo de similaridade.
O Método CNN destacou-se por analisar as características das imagens, incluindo cores, texturas e formas, permitindo a busca e recuperação de imagens. A coleta das informações permitiu a indexação, geração e utilização de imagens para redes neurais e/ou para uso no Método CBIR, implementado por meio de algoritmos de análise de histograma e comparação de imagens.

Conclusões

Os resultados demonstraram uma excelente acurácia de 0.9984 e baixa perda no conjunto de treinamento, bem como uma perda e precisão razoáveis na validação, indicando que o modelo aprendeu efetivamente.
Houve desafios iniciais de acurácia, com aprimoramento do método, com ajustes de parâmetros e otimização do cálculo de similaridade.
Para pesquisas futuras, sugere-se expandir o conjunto de dados, explorar técnicas avançadas de *deep learning*, contribuindo para a agricultura de precisão e o diagnóstico de doenças de plantas de forma mais eficaz.

Agradecimentos

O autor agradece pelo suporte fornecido pela instituição de pesquisa e pela orientação ao longo deste estudo.

Referências

KHAN, Asifullah et al. **A Survey of the Recent Architectures of Deep Convolutional Neural Networks. Artificial Intelligence Review**. v. 53. p. 5455-5516. 2020

Keras. Disponível em: https://keras.io .Acesso em 12 abr 2023

TensorFlow. Disponível em: https://www.tensorflow.org/learn. Acesso em: 12 abr. 2023.