|  |  |  |  |
| --- | --- | --- | --- |
| Title | Self-supervised clustering for leaf disease identification | | |
| Author(s) Name | Muhammad Mostafa Monowar, Md Abdul Hamid, Faris A Kateb, Abu Quwsar Ohi, MF Mridha | | |
| Contact Email(s) | firoz.mridha@aiub.edu | | |
| Published Journal Name | Agriculture | | |
| Type of Publication | Journal | | |
| Volume | 12 | Issue | 6 |
| Publisher | MDPI | | |
| Publication Date | 2022/6/5 | | |
| ISSN | 2227-7390 | | |
| DOI | 10.3390/agriculture12060814 | | |
| URL | https://www.mdpi.com/2077-0472/12/6/814 | | |
| Other Related Info. |  | | |
|  | | | |

|  |  |
| --- | --- |
| Abstract |  |
| Plant diseases have been one of the most threatening scenarios to farmers. Although most plant diseases can be identified by observing leaves, it often requires human expertise. The recent improvements in computer vision have led to introduce disease classification systems through observing leaf images. Nevertheless, most disease classification systems are specific to diseases and plants, limiting method’s usability. The methods are also costly as they require vast labeled data, which can only be done by experts. This paper introduces a self-supervised leaf disease clustering system that can be used for classifying plant diseases. As self-supervision does not require labeled data, the proposed method can be inexpensive and can be implemented for most types of plants. The method implements a siamese deep convolutional neural network (DCNN) for generating clusterable embeddings from leaf images. The training strategy of the embedding network is conducted using AutoEmbedder approach with randomly augmented image pairs. The self-supervised embedding model training involves three different data pair linkage scenarios: can-link, cannot-link, and may-link pairs. The embeddings are further clustered using k-means algorithm in the final classification stage. The experiment is conducted to individually classify diseases of eight different fruit leaves. The results indicate that the proposed leaf disease identification method performs better than the existing self-supervised clustering systems. The paper indicates that end-to-end siamese networks can outperform well-designed sequentially trained self-supervised methods. | |