

Title: Automated classification of diseased cauliflower: a feature-driven machine learning approach

Author Names: Mala Rani Barman, Al Amin Biswas, Marjia Sultana, Aditya Rajbongshi, Md. Sabab Zulfiker, Tasnim Tabassum

Abstract: Cauliflower is a popular winter crop in Bangladesh. However, cauliflower plants are vulnerable to several diseases that can reduce the cauliflowers' productivity and degrade their quality. The manual monitoring of these diseases takes a lot of effort and time. Therefore, automatic classification of the diseased cauliflower through computer vision techniques is essential. This study has retrieved ten different statistical and gray-level co-occurrence matrix (GLCM)-based features from the cauliflower image dataset by implementing a variety of image processing techniques. Afterwards, the SelectKBest method with the analysis of variance f-value (ANOVA F-value) has been used to identify the most important attributes for classification of the diseased cauliflower. Based on the ANOVA F-value, the top N ($5 \leq N \leq 9$) most dominant attributes is used to train and test five machine learning (ML) models for classification of diseased cauliflower. Finally, different performance metrics have been used for evaluating the effectiveness of the employed ML models. The bagging classifier achieved the highest accuracy of 82.35%. Moreover, this model has outperformed other ML classifiers in terms of other performance metrics also.

Keywords: Analysis of variance Bagging classifier Cauliflower Image processing Machine learning

DOI: <http://doi.org/10.12928/telkomnika.v22i4.25812>

<http://telkomnika.uad.ac.id/index.php/TELKOMNIKA/article/view/25812>