TABLE III
VEGETATION INDICES AND THE CORRESPONDING APPLICATION FOR PLANT DISEASE DETECTION

| Vegetation indices | Equation | Application areas | References | References for disease detection using vegetation index |
|---|---|---|---|---|
| Photochemical Reflectance Index (PRI) | (R570−R531)/(R570 + R531) | Yellow rust of wheat due to photosynthesis and carotenoid content | [238] | [239–241] |
| Anthocyanin Reflectance Index (ARI) | 1/R510-1/R700 | Yellow rust of wheat due to anthocyanin content | [192] | [167, 242–244] |
| Structure Intensive Pigment Index (SIPI) | (R800− R445)/R800−R680) | Ratio of carotenoids and chlorophyll | [245] | [243, 246] |
| Water Index (WI) | R900/R970 | Evaluates the water stress | [247] | [248–250] |
| Normalized Difference Vegetation Index (NDVI) | (R800−R670)/(R800 + R670) | Estimate vegetation coverage based on chlorophyll content | [251] | [252–254] |
| Green Normalized Difference Vegetation Index (GNDVI) | (NIR− GREEN)/(NIR + GREEN) | Withered/ aged crops, also measure nitrogen content in leaves | [255] | [253, 254, 256] |
| Ratio Vegetation Index (RVI) | R800/R670 | Estimates the plant biomass | [257] | [258, 259] |
| Carotenoid Reflectance Index (CARI) | 1/R510−1/R550 | Estimates carotenoid content | [260] | [261–263] |

matrix is developed, several texture descriptors, such as contrast, energy, correlation, and homogeneity, can be evaluated. These features provide information on the gray-level spatial features for the images [218]. Besides, the Gabor filter, which is used in texture analysis in land cover classification, can detect texture variations in plants due to diseases [219], [220]. Also, the Fourier transform, which is used to identify recurring patterns in land cover classification tasks, can be used to obtain the patterns in spectral signatures and detect healthy and infected plants [221], [222]. Hence, exploring and integrating these techniques could provide better features to process the hyperspectral data.

Feature extraction is vital as it defines how well the model can classify the disease types. During the feature extraction process, new plant disease feature vectors are formed, which combine the spatial and spectral features and then feed into classification-based models. In some instances, vegetation indices can be considered the features to feed onto the models [17]. Vegetation indices are obtained by combining spectral features of multispectral satellite images at the VNIR range of the green plants [223]. However, for hyperspectral images, vegetation indices are defined at the plants' medium and NIR bands to detect the diseases. The vegetation indices simplify the processing of the hyperspectral images as they focus on particular bands where the conditions or changes are observed. This helps to reduce the processing cost of hyperspectral images as well [176].

There are more than 40 types of vegetation indices that can be used for changed environmental conditions [224], crop yield estimation [225], vegetation classification [224], and host-pathogen interactions [176]. A study in [226] analyzed the relationship between vegetation indices and early disease symptoms of grapevines. The study showed that vegetation indices could help detect grapevines' plant disease (tiger-stripe) and identify the possible reason for the fungal diseases. Another study [158] showed that using vegetation indices and an SVM classification model, the diseased leaves of sugar beets could be identified with an accuracy of 97%. Table III provides some of the vegetation indices used for different plant disease detection types.

Although vegetation indices provide good results, they take the hyperspectral images' full spectral information. Therefore, a combination of other techniques such as PCA, linear discriminant analysis (LDA), and minimum noise fraction techniques (MNF) [17] can be used to improve the feature extraction process. PCA can capture the important features in the spectral data and reduce the dimensionality of the data. The LDA can then be employed to improve the separability of the diseased and healthy datasets in the feature space. MNF can help to discard the redundant noisy spectral information and focus on the important spectral band information. The combination of feature extraction algorithms improves the overall performance of the models, as it can now focus on the most informative aspects of the data [227].

*3) Modeling and Data Interpretation:* The final step of the hyperspectral image analysis is to choose a model based on the target/objective of the research. For instance, regression or classification models can be employed for plant disease diagnosis/detection. Regression models aim to predict the targeted variable (such as disease), and the classification models aim to distinguish target variables [17]. Generally, two types of models are used to perform the hyperspectral images: 1) machine-learning-based classification models, which include neural network [254], SVM [158], random forest (RF) [255], k-nearest neighbor (kNN) [256], spectral angle mapper (SAM) [257], and maximum likelihood algorithms [258]; and 2) regression-based models, which include partial least squares regression (PLSR) [259], Dirichlet aggregation regression [7], logistic regression [260], and multiple linear regression [261]. Neural networks are machine-learning-based algorithms consisting of interconnected nodes that learn the pattern and perform the prediction. The neural network is trained on the labeled images/data of the plants. Once the training is completed, the neural network can be used for predicting the health condition of the plants in the new datasets or unlabeled datasets [3], [262]. SVM is a supervised machine learning algorithm that helps to find the best hyperplane that separates data points of different classes. For plant disease detection, it can be used to differentiate between healthy and diseased plant samples based on their

Authorized licensed use limited to: FUDAN UNIVERSITY. Downloaded on August 28,2025 at 01:49:24 UTC from IEEE Xplore. Restrictions apply.