area and the research objectives. Once the correct options are chosen and images are obtained, the next step is to preprocess the images. The preprocessing of the hyperspectral images involves atmospheric compensation and calibration. The hyperspectral image data obtained from the airborne platforms can be distorted by the Earth's atmosphere due to the scattering of molecules and absorption of gases such as oxygen, carbon dioxide, and methane. Hence, it is paramount to adjust the raw hyperspectral images to avoid the impact of atmospheric compensation [185]. The empirical line method (ELM) is commonly employed to mitigate the atmospheric compensation method, which converts the measured radiance data to surface reflectance values to obtain more insights into the surface materials [186]. It requires the presence of calibrated panels and known spectral reflectance within the scene. However, ELM-based approaches are only suitable for airborne or UAV-based data acquisition platforms but not for satellite or large-scale remote sensing and monitoring [186]. Depending on the dataset, radiative transfer modeling and some hybrid approaches [186] can also correct atmospheric compensation.

On the other hand, calibration ensures the standardization of the spatial and spectral dimensions, eliminates the curvature effects, and adjusts the instrument errors so that the information remains intact under different working conditions [187], [188]. One standard practice is to employ reflection calibration [17]. The reflection calibration employs one black and one white image as a reference. The black image is obtained by covering the camera lens with an opaque cap. The white image is obtained using high reflectance materials such as a Teflon white surface board or a Spectralon panel [189] to acquire the highest intensity for each pixel corresponding to each wavelength. The black and white images are then used as references to calibrate the acquired hyperspectral images. Equation (1) shows a basic and simplified mathematical equation for the image calibration of hyperspectral images. However, this equation has a similarity to the ELM

$$R = \frac{R_I - B_I}{W_I - B_I} \tag{1}$$

where R is the calibrated/corrected image, R_I is the raw image, B_I is the black image, and W_I is the white image.

The surface curvature effects are adjusted through adaptive spherical transform [190], Lambert transform [191], or with image normalization techniques [192]. Spectrum adjustment is performed through smoothing algorithms such as median filter, moving average, Gaussian filters, wavelet, and Fourier transforms [17] to enhance the image quality and reduce the noise in spectral images. Apart from the abovementioned techniques, standard normal variate and multiplicative scattering correction are also employed to minimize the spectral variability due to the spectrum scattering [17].

In recent years, computer vision techniques, especially machine learning and deep learning techniques, have revolutionized the traditional image preprocessing methods. A recent study in [193] developed an end-to-end deep learning technique using residual, parallel-scale, and multiscale networks to reconstruct RGB images from the raw hyperspectral images. The model reduced the cumulative errors and the computational involution.

Another study proposed to use image stitching techniques [194] by using Pix4Dmapper [195] software to adjust the resolution of remote sensing images.

2) Feature Extraction: Once the calibration and adjustment of the images are completed, the next step is to extract useful information from the hyperspectral images. One of the most popular methods to extract information is using image segmentation. Image segmentation aims to extract the features of the target element (for instance, the disease types) from the background and assigns masks for each target for further processing [17]. The most popular segmentation techniques that are reported in the literature for plant disease detection through hyperspectral images are K-means [41], water-shed [196], edge detection [197], [198], and threshold-based algorithms [12]. K-means algorithms can isolate the healthy plant regions from the disease regions through their spectral profile. They are relatively simple to implement and can be used to process larger datasets [199], [200]. Watershed algorithms perform segmentation on the hyperspectral images by using the topographical gradient. This algorithm helps to segment the area of the interest (infected/diseased) region from the rest of the healthy images. Thus, they are suitable for analyzing complex images with multiple diseases present [201]. Edge detection algorithms estimate the edges of the defected region, which helps to differentiate between the diseased and healthy regions. This can help to detect the diseases early [202], [203]. Threshold-based algorithms classify the pixel based on the hyperspectral spectral intensity variations to distinguish between healthy and diseased plants [12], [204]. Besides image segmentation, PCA [205], [206], [207], wavelet transform [208], [209], [210], spectral indices [211], [212], partial least squares (PLS) [213], and neural network architectures are also used [214], [215], [216] for feature extraction. PCA is a well-known statistical method that reduces the dimensionality of the hyperspectral data into a new coordinate system so that the variances are preserved [205], [206], [207]. Wavelet transform helps to obtain a multiresolution analysis of the hyperspectral data, which provides the features from different scales [208], [209], [210]. Spectral indices combine the reflectance values measured at the different spectral wavebands of hyperspectral data. It helps to understand and extract contrast information of the image data [211], [212]. PLS is also a multivariate regression method, which reduces the dimensionality of the hyperspectral image dataset by extracting the latent variables [213]. Finally, neural networks are machine-learning-based algorithms. The neural network extracts features from hyperspectral image data using input information such as spectral band. Then, the information is processed through one or multiple layers of neurons. Each neuron extracts and learns the complex features, which allows us to know the spectral profile (features in the shallow layer) and disease profile (features in the deeper layer) [214], [215], [216].

In addition to the abovementioned feature extraction techniques, the gray-level co-occurrence matrix (GLCM), a commonly used spatial feature extraction technique in land cover classification, can be employed to detect plant diseases [217]. At first, spatial information direction/distance is chosen to build a co-occurrence matrix and symmetrized. Once the GLCM