CHAPTER II
RELATED CONCEPTS AND LITERATURE REVIEWS

This section uses UAV images to explain concepts and theories for detecting weeds in crop fields. First, the importance of cassava and the processes for growing cassava were summarized in the topic of cassava. Details of RGB data and their indices are described in the spectral discrimination of vegetation because these are crucial components for weed discrimination. Finally, critical theories for developing the classifier were summarized, including image segmentation, mean-shift filtering, and K-means clustering.

2.1 Cassava

2.1.1 Important of cassava

Cassava or tapioca (*Manihot esculenta Crantz*) is one of the most important economic crops in Thailand. Products of cassava are fresh root, starch, chip, and pellets. Most of them are raw materials in food, pharmaceutical, animal feed, sweeteners, textiles, monosodium glutamate (MSG), and ethanol. Thailand is the third-largest cassava producer and the world's largest exporter of cassava products (FAOSTAT Statistical Database, 2022). Approximately 8 million tons of cassava were exported in 2018, valued at around 74 billion bahts (Arthey, Srisompun, and Zimmer, 2018; Office of Agricultural Economics, 2021).

In Thai agriculture, cassava has increased in the harvested area since the 1970s (Figure 2.1), covering more than 1.7 million hectares in 2021 (FAOSTAT Statistical Database, 2022; OAE, 2022). Figure 2.2 presents Thailand’s cassava production of 2020 at the provincial level, with most of the cassava produced in the Northeastern region. The five provinces that produced the highest cassava production included Nakhon Ratchasima (4.6 million tons), Kamphaeng Phet (2.3 million tons), Chaiyaphum (1.8 million tons), Kanchanaburi (1.6 million tons), and Ubon Ratchathani

(1.5 million tons). Cassavas are called insurance crops because the farmer can harvest and sell them at any time, not only during the period as sugarcane (Polthanee, 2018). Thus, cassavas are an essential source of cash income for farmers.

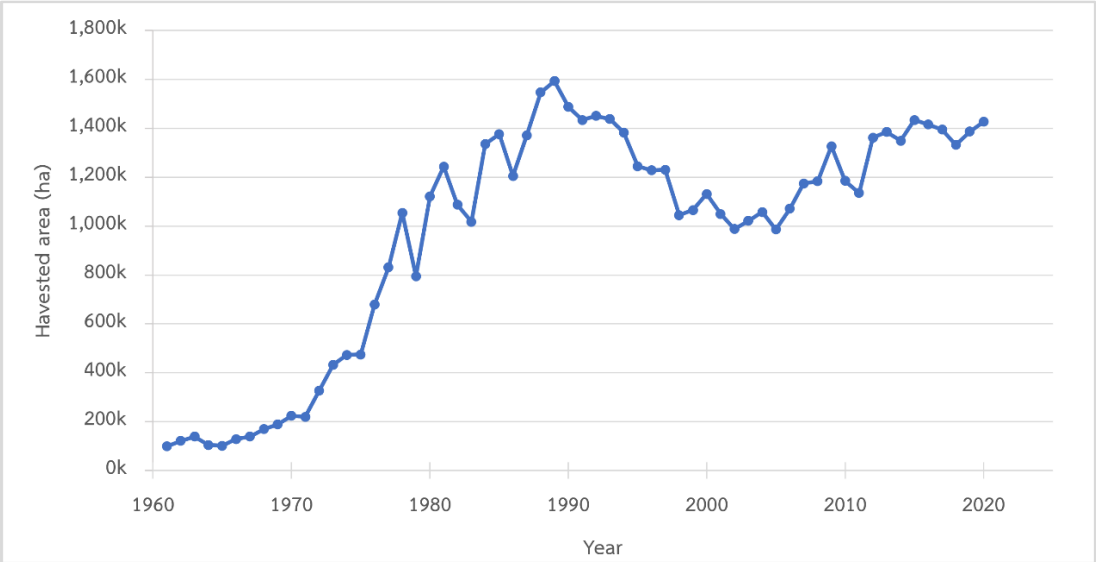


Figure 2.1 Cassava harvested area of Thailand.

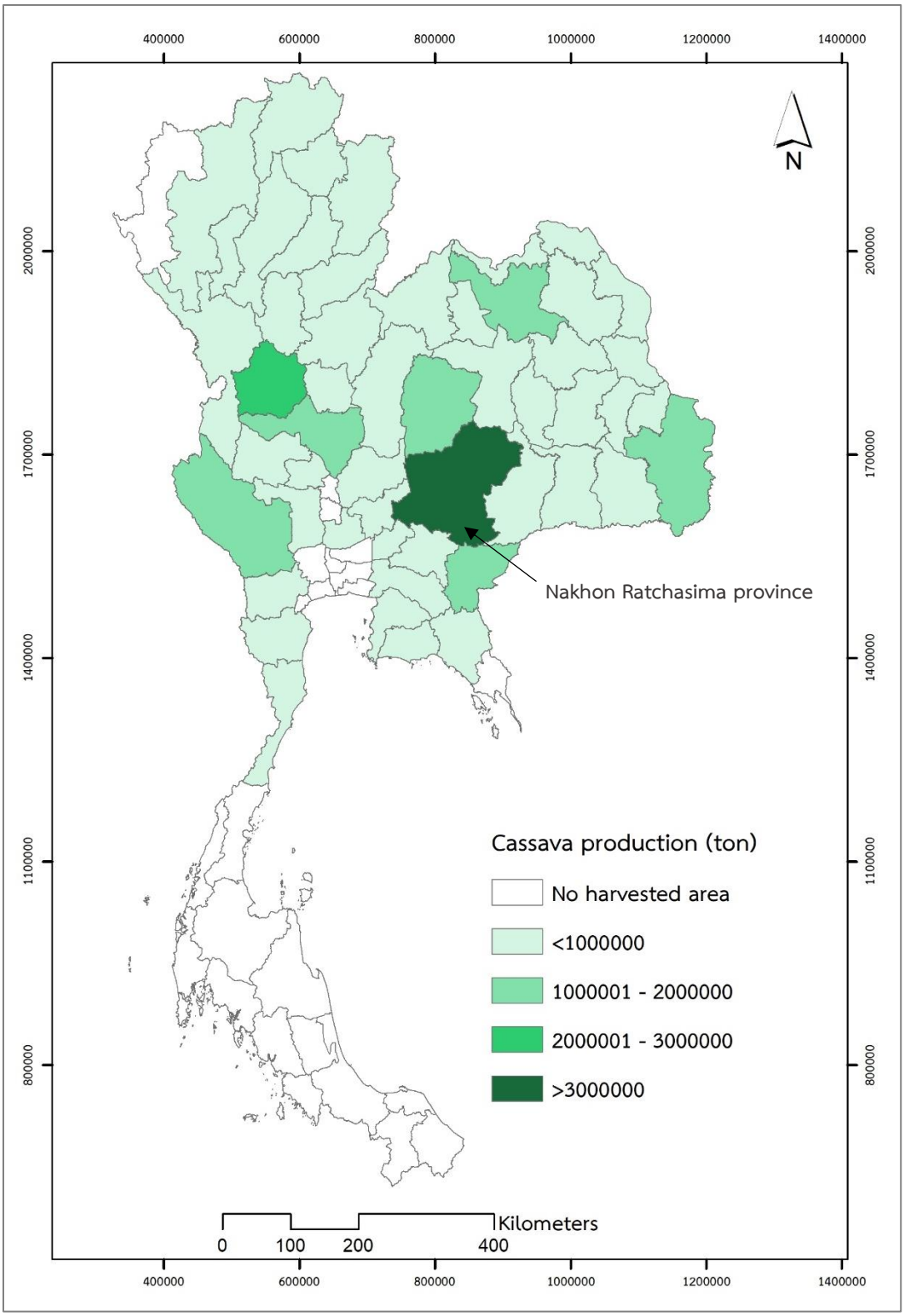
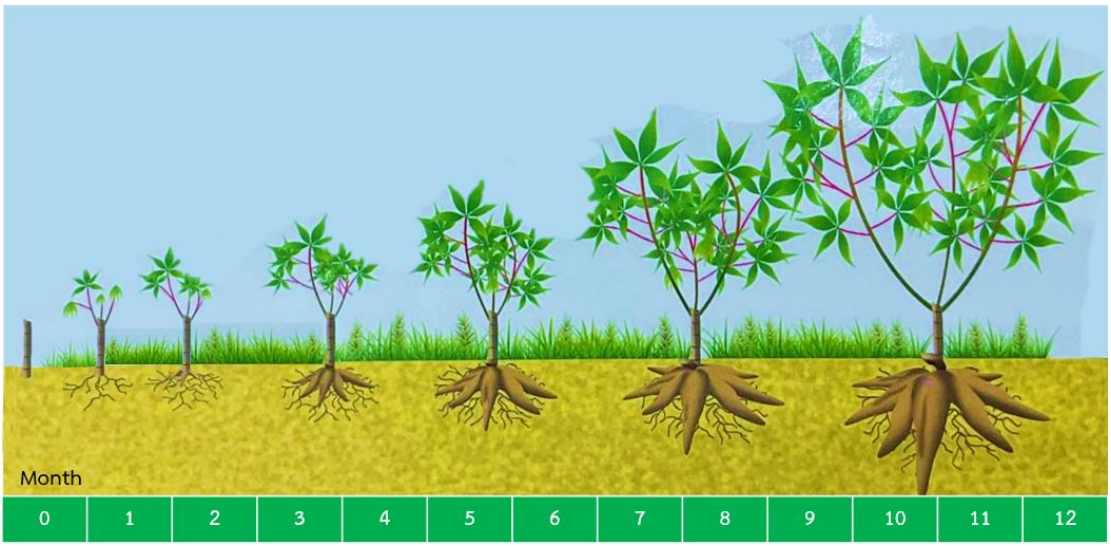


Figure 2.2 Cassava production in Thailand.

2.1.2 Growing cassava

Cassava is a drought-tolerant crop grown in dried soils and adapted well to various cultivating conditions (Howeler et al., 2013). Although cassava is simple to grow, weeds and diseases have dominantly affected the productivity of cassava yield (Wydra and Verdier, 2002). Cassava requires a long growing period; growth states are shown in Figure 2.3.



Source: Plant protection research and development office (2016).

Figure 2.3 Cassava growth stages.

Typically, Thai farmers plant cassava before the rainy season (March-April) or after the rainy season (November-December). The processes of cassava planting include cassava variety preparation, land preparation, pest control, and harvesting. Details of the processes are as follows:

(1) Cassava varieties and preparing

There are three groups of cassava planting in Thailand: sweet type (for direct consumption), bitter type (for processing), and ornamental type. Most cassavas planted in Thailand are bitter, have high starch content, and are usually used as raw industrial materials. The cassava species included Rayong 1, Rayong 3, Rayong 5, Rayong 7, Rayong 9, Rayong 11, Rayong 60, Rayong 72, Rayong 90, Kasetsart 50, Huay Bong 60, Huay Bong 80, Sriracha 1 and others (OAE, 2022).

Cassava is planted with stems cut about 20-35 centimeters from the middle of a healthy and free of deceases and pathogens cassava. Farmers always soak the stems in pesticides before planting to eliminate diseases and pathogens.

(2) Land preparing

Cassava requires a loose texture and well-drained soil to facilitate the growth of roots and clearing from weeds to prevent competition with young crops in its initial state. The land should dry for at least two weeks and till again to control weeds. The land should form into ridges (Figure 2.4) with a space of 0.8-1 meter and a height of 30 centimeters. Cassava stems are usually planted at a distance of 0.8-1 meter between and within a row.

Source: Department of Agriculture (2001).

Figure 2.4 Land preparation into ridges forms for cassava planting.

(3) Pest management

Cassavas are weak and require well-managed in the first four months after planting or until the cassava canopy is closed. Pest management is done according to the age range, as shown in Table 2.1.

Table 2.1 Pest management for cassava.

Age (month)	Pest	Management
0	Mealy bug and Scale insect	Soak stems in pesticides
	Stem bright disease	Soak stems in fungicide
	Weeds	Apply pre-emergent weed control
1	Witches' broom and bright stem disease	Remove stem and burn outside the fields
2-3	Red mite	Apply insecticide
	Mealy bug	Control by insects (parasitic wasps or green lacewing) or apply insecticide
	Anthracnose, root rot, and root-knot disease	Apply fungicide
	Weeds	Apply post-emergent herbicide or remove them with human labor
4	Anthracnose, root rot, and root-knot disease	Apply fungicide
	Witches' broom disease	Remove the disease stem and burn outside the fields, and apply fungicides to the rest of the crops
	Weeds	Apply post-emergent herbicide or remove them with human labor

(4) Harvesting

Cassava roots were harvested 8-14 months after being planted, and around 70 percent of cassavas in Thailand were harvested at 10-12 months (OAE, 2022). Cassavas were harvested by cutting the stem above the ground at least 30 centimeters and digging the roots out by hand or machine. Due to the growing conditions, average cassava yields are 21-23 tons per hectare (Arthey et al., 2018).

2.1.3 Weed control

Weeds are a barrier to growing crops, which causes the loss of quality and quantity of crops. Not only snatching up the essential nutrients of crops, but weeds are also a habitat of pests, including aphids, mealybugs, scarlet mites, and phytoplasma, the causative agent of cassava bush disease (Plant protection research and development office, 2016). Because cassavas are slowly grown in the initial period, weeds have a high opportunity to compete with the young cassavas. An extreme invasion of weeds can cease the growth and reduce up to 80% of cassava yields (Jeamjamnanja et al., 1984).

Controlling weeds is an essential process in cassava growing. The appropriate time for controlling weeds is in the early period before the canopy closes, at 3, 8, and 12 weeks after planting (Ekanayake, Osiru, and Porto, 1997). For the first 3-4 months, cassava leaves, stems, roots, and tubers are produced. Several tubers developed in this period; thus, this was crucial for cassava. However, cassava cannot compete well with weeds in this period because the root system is young and weak. Weed competition, especially the fast-growing weeds, is obstructed for canopy and tubers development. After planting for four months, the tuber size is developed, the cassava canopy grows and connects to the neighbors, the canopy will shade, and most weeds cannot compete with cassava.

Common approaches for controlling weeds in crops are physical control (removal by mowing, mulching, tilling, burning, grazing, or by hand), cultural practices (tillage, irrigation), and chemical control (use of chemical substances). Herbicides, chemical substances used for weed control, are often used to control weed in a large field or when lacking labor. Weed controlling costs one-third of the total cost of cassava growing (Plant protection research and development office, 2016), which is trending to increase due to the increasing price of herbicides (McGeeney, 2022).

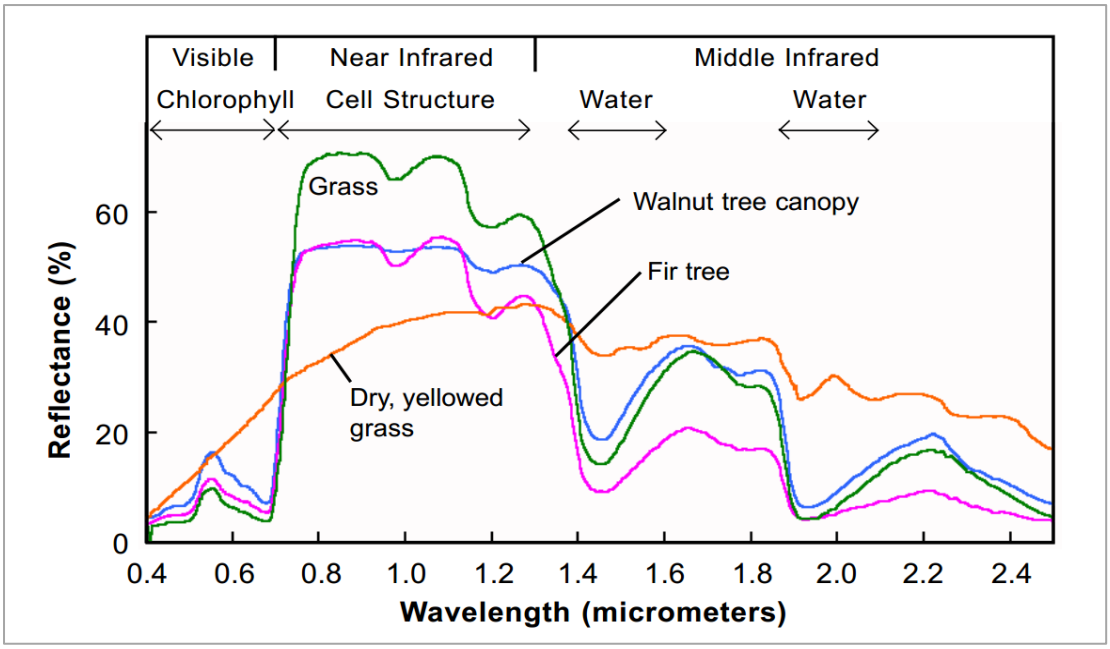
Herbicides for weed control in crops are applied by spray equipment using human labor, which has some limitations, including difficulty accessing the complex area with the terrain and managing the number of herbicides, and the effect of herbicides on human health has been a concern. Recently, UAV spray technology has been developed to serve agricultural purposes, including pesticide and herbicide

applications. Regarding the combined advantages of UAV spraying and weed identification systems, herbicide control becomes more cost-effective and safer for farmers' health.

2.2 Spectral discrimination of vegetation

The relation of electromagnetic radiation with plants varies with the radiation wavelength. Each vegetation type has a signature spectral pattern that can identify its species. Also, the same plant leaves will exhibit significant differences in how they reflect light depending on health and vigor (Woolley, 1971). The amount of radiation reflected from plants is inversely related to radiation absorbed by plant pigments and varies with the wavelength of incident radiation (Mulla, 2013). The reflectance of vegetation in various spectrum wavelengths is shown in Figure 2.5.

The study of vegetation is mainly based on near infrared (NIR) wavelength because it has a strong reflectance, as seen in Figure 2.5, which can clearly show the differences in plants. However, a sensor for collecting the NIR data was more costly and more difficult to mount to UAVs because it is not available with optical sensors. Moreover, the NIR-based images were difficult to verify the anomaly of vegetation without comparing it to natural color (G. E. Meyer and Neto, 2008). According to the visible spectral, the reflectance of plants is shared the same pattern but is different in percent of reflectance due to the different colors and shades of leaves (Gates, Keegan, Schleter, and Weidner, 1965), which refer to plant species and health of the plant. These properties make the optical sensors able to separate objects also plant types. The studies of Bah, Hafiane, and Canals (2018); Gao et al. (2018); Gašparović et al. (2020); Huasheng Huang et al. (2018); Lottes et al. (2017); Subeesh et al. (2022); Woebbecke, Meyer, Bargen, and Mortensen (1995) showed that the low-cost RGB sensor was applied for weed identification.



Source: Smith (2012).

Figure 2.5 Reflectance spectra of different types of green vegetation compared to a spectral signature for senescent leaves.

RGB-color images with red (R), green (G), and blue (B) channels have been used to detect information in plants such as stress, disease, and type (Chaudhary, Chaudhari, and Godara, 2012; Hall et al., 2018; Lameski, Zdravevski, Trajkovik, and Kulakov, 2017; Zakaluk, Sri Ranjan, and Ranjan, 2008). The data provided from UAV images are DN values, while the signature of objects is presented by reflectance spectra value. The DN data can represent the trend of the reflectance value. Thus, the indices calculated from the DN data also provided qualities of the object's spectral signature (Candiago, Remondino, De Giglio, Dubbini, and Gattelli, 2015).

The index is a spectral transformation of two or more bands designed to enhance the contribution of vegetation properties. Indices have been widely used for evaluating vegetation cover, growth dynamics, and biomass (Jinru and Su, 2017). Using color indices in crop and weed detection has various advantages. For example, the indices highlighted green color from the images, less sensitive in light conditions and can extract information which challenging to distinguish and compare by humans (G. E. Meyer and Neto, 2008).

The UAV image quality is susceptible to illumination; rapid changes in weather conditions may cause non-smooth light in the images. Several studies in weed detection using UAV images have developed the processes of image processing techniques to solve illumination problems. Illumination issues play an essential role in the classification process. Color indices have been proven to reduce illumination problems in normal light conditions (Hamuda, Glavin, and Jones, 2016). However, the indices vary depending on crop types and intense light conditions (Hamuda et al., 2016). The widely used RGB-based indices showing plant and soil properties are shown in Table 2.2.

Table 2.2 Spectral indices based on RGB values to distinguish vegetation and non-vegetation.

Index	Equation	Dominant	Reference
Excess Red (ExR)	$1.4 \times r - g$	red spectrum extraction	G. Meyer, Hindman, and Laksmi (1999)
Excess Green (ExG)	$2g - r - b$	green spectrum extraction	Woebbecke et al. (1995).
Excess Blue (ExB)	$1.4 \times b - g$	blue spectrum extraction	Guijarro et al. (2011)
Excess Green minus Excess Red (ExGR)	$ExG - ExR$	highlight vegetation	G. E. Meyer and Neto (2008)
Normalized Green Red Difference Index (NGRDI)	$\frac{G - R}{G + R}$	vegetation discrimination	Tucker (1979)
Green Leaf Index (GLI)	$\frac{2G - R - B}{2G + R + B}$	vegetation discrimination	Louhaichi, Borman, and Johnson (2001)
Visual Atmospheric Resistance Index (VARI)	$\frac{G - R}{G + R - B}$	vegetation discrimination	Gitelson, Kaufman, Stark, and Rundquist (2002)

Table 2.2 Spectral indices based on RGB values to distinguish vegetation and non-vegetation (Continued).

Index	Equation	Dominant	Reference
Brightness Index (BI)	$\sqrt{\frac{G^2+R^2}{2}}$	soil discrimination	Mathieu, Pouget, Cervelle, and Escadafal (1998)
Color Index (CI)	$\frac{R-G}{R+G}$	soil discrimination	Escadafal and Huete (1991)

where:

R, G, B is digital number of red, green, and blue channels (0-255)

$r = \frac{R^*}{R^*+G^*+B^*}, g = \frac{G^*}{R^*+G^*+B^*}, b = \frac{B^*}{R^*+G^*+B^*}$

$R^* = \frac{R}{255}, G^* = \frac{G}{255}, B^* = \frac{B}{255}$

2.3 Image segmentation

Image segmentation is a process for segmenting or grouping the objects in the images. Segmentation is an essential process for image processing applications, which is applied in many fields, such as medicine, object recognition, and agriculture. In the agricultural field, the segmentation process is mainly used to segment plants from the background (soil and sky) and segment plant types (crop and weed). Segmentation is an important process in image analysis since the segmentation results impact the analysis's quality and performance in automatic processing (Morales et al., 2011).

There are various ways to categorize image segmentation approaches, for example, thresholding, edge-based, region-based, clustering, watershed, Partial Differential Equation (PDE), Artificial Neural Network (ANN), and neural network. The main approaches were summarized as follows:

(1) Edge Based

Segmentation based on edge detection is a method for locating the boundary of objects. The edges or boundaries were identified from the rapid change of intensity or pixel brightness value, then connected to form closed object boundaries. The edge location was detected by following one of these criteria (Kaur and Kaur, 2014; Narkhede, 2013):

1) the first derivative of intensity is larger than the threshold, or

2) the second derivative of intensity is zero crossing.

This approach was sensitive to noise. Thus, it requires an image enhancement process to reduce or remove noise (Kuruvilla, Sukumaran, Sankar, and Joy, 2016; Narkhede, 2013). There are many techniques for detecting edges, for example, Sobel, Laplacian, Canny, and Laplacian of Gaussian detection.

(2) Threshold-based

Image segmentation based on thresholding is the simplest method for segmentation. The threshold values are calculated from a binary image, which can be manually defined or automatically calculated from the image. The threshold value is the criteria for separating pixels of an image into the region or separating objects from the background. This approach is suitable for segmenting the light object on a shaded background but is sensitive to light illumination (Kuruvilla et al., 2016).

(3) Region-based

Segmentation based on region is a process of partitioning an image into a region. Regions are used to represent the whole object or some part of the object in the image (Kuruvilla et al., 2016). The region-based segmentation used value similarity (which contains grey value differences and variance) and spatial proximity (which contains Euclidean distance and compactness of a region) for partitioning pixels (Narkhede, 2013). This approach also requires the thresholding technique. There are two main types of region-based segmentation: region growing and splitting and merging. Some researchers grouped the watershed method and clustering method in the region-based segmentation approach because these methods produced

segmentation based on the similarity of pixel data (Raja, Abdul KhadiRr, and Ahamed, 2009).

(4) Cluster-based

Image segmentation based on clustering is a method for grouping objects using similarity properties. Pixels with similar characteristics were grouped together to form clusters. There are two main types of clusters: hard clustering and soft clustering. A pixel belongs to only one cluster in the hard, complex clustering type. An example of a hard clustering type is K-means clustering. While soft clustering is more flexible, a pixel is partitioned into a cluster; thus, one pixel can belong to more than one cluster. Fuzzy C-means is an example of a soft clustering type.

Segmentation is an essential process for digital image processing, especially high-resolution images. This process aims to remove the complexities of the image and or enhance or edit the image's appearance. However, implementing each method is suitable for different images and purposes of use. There are some advantages and disadvantages of each segmentation approach summarized in Table 2.3.

Table 2.3 Advantages and disadvantages of the segmentation approach.

Approaches	Advantages	Disadvantages
Edge-based	• Results are similar to the human perspective • Suit for image with good contrast between objects	• Not suited for images with ill-defined edges or many edges • Sensitive to noise • Difficult to produce close curve boundary
Threshold-based	• Fast processing • Noncomplex for computational • Not require prior information	• Sensitive to noise • Depending on histogram peaks • Not suited for images with unapparent peak (broad or flat)

Table 2.3 Advantages and disadvantages of the segmentation approach (Continued).

Approaches	Advantages	Disadvantages
Region-based	• Suit for image with good contrast between objects • Work well with the homogeneous region • Consider spatial information • Provide region continuity result	• Require more time and memory for computational • Not suited for images with noise
Cluster-based	• Fast processing for a small number of clusters • Eliminate noise • Can process automatically	• Difficult for the high-value clusters

These segmentation approaches are based on an iterative technic with stop criteria limitations. The stop criteria require much memory or involve threshold parameters that are difficult to optimize (Tian, Hsiao-Chun, and Huang, 2014). Mean-shift filtering is a non-parametric method. This method uses entropy as a stop criterion for the iterative process. The entropy of the region or group was reduced in an iterative process until it reached a stable value (convergence) when the image became homogeneous. The study of Rodríguez et al. (2011) proved that the results from mean-shift filtering were similar to results obtained by segmentation; the image presented natural aspects and still preserved details of the original images. Dorin Comaniciu and Meer (2002) and Grenier, Revol-Muller, and Gimenez (2006) also state that mean-shift filtering gives products close to products from clustering-based segmentation. Thus, segmentation can be obtained directly through filtering (Morales et al., 2011; Rodríguez et al., 2011).

2.4 Mean shift filtering

Mean shift is a non-parametric iterative algorithm developed by Fukunaga and Hostetler (Fukunaga and Hostetler, 1975) and has been widely used for filtering and clustering analysis in computer vision and image processing. The mean shift shows a high potential for edge preservation in filtering and segmentation processes (D. Comaniciu and Meer, 1999). The mean-shift method considers the image's geometric and photometric during the process (Song, Gu, Cao, and Viberg, 2006). The theory of the mean-shift algorithm and mean-shift filtering summarized by Li (2012) are described below.

The mean-shift algorithm can be considered a gradient ascent on the density function. This problem can be written as:

$$\nabla \hat{f}(x) = 0 \tag{1}$$

Using a non-parametric kernel-based probability density function estimation to solve the problem in equation (1), the quantity m(x) is called the mean-shift vector as the following:

$$m(x) = x^{[t+1]} - x^{[t]} = \frac{\sum_{i=1}^{n} g\left(d^2\left(x^{[t]}, x_i, H\right)\right).x_i}{\sum_{i=1}^{n} g\left(d^2(x^{[t]}, x_i, H)\right)} - x^{[t]} \tag{2}$$

Where t is the iteration variable, d is Mahalanobis distance, $x^{[0]}$ is set to a given x_i, and $g(.)$ is a weight function deriving from the kernel function.

The procedures of mean-shift filtering are described as follows. Let $x = (x_s, x_r)$ be vector and bandwidth matrix H composed of $[H_s, H_r]$. Where subscript s and p represent spatial and range (pixel value) of vector and bandwidth matrix. The bandwidth matrix can be defined as

$$H = \begin{bmatrix} H_s & 0 \\ 0 & H_r \end{bmatrix} = \begin{bmatrix} {h_s}^2 \cdot I^p & 0 \\ 0 & {h_r}^2 \cdot I^q \end{bmatrix} \tag{3}$$

where h_s is the scalar value of spatial bandwidth, h_r is the scalar value of range bandwidth, I is the identity matrix, p is the number of spatial dimensions, and q is the

number of range dimensions of an image. For the three-layer image, p and q are 2 and 3, respectively.

The kernel in one-dimensional is expressed as:

$$K(x) = [x_s, x_r] = \frac{c}{(h_s)^p \cdot (h_r)^q} K\left(\left\|\frac{x_s}{h_s}\right\|^2\right) \cdot K\left(\left\|\frac{x_r}{h_r}\right\|^2\right) \quad (4)$$

A color image that includes three feature space components is taken to illustrate the procedures of mean-shift filtering. Let $\{x_i\}_{i=1,2,\ldots,n}$ and $\{z_i\}_{i=1,2,\ldots,n}$ be the original and the filtered image points in the d-dimensional feature space, respectively. The mean-shift filtering algorithm can be summarized as:

(1) Initialize $x^{[0]}$ to a given x_i,

(2) Initialize a bandwidth matrix $H = \begin{bmatrix} {h_s}^2 & 0 & 0 & 0 & 0 \\ 0 & {h_s}^2 & 0 & 0 & 0 \\ 0 & 0 & {h_r}^2 & 0 & 0 \\ 0 & 0 & 0 & {h_r}^2 & 0 \\ 0 & 0 & 0 & 0 & {h_r}^2 \end{bmatrix}$,

(3) Compute $x^{[t+1]}$ the following equation (2) until convergence,

(4) Assign to z_i the spatial position of x_i and the range value of $x^{[t+1]}$, so $z_i = ((x_i)_s, (x^{[t+1]})_r)$,

(5) Iterate the above steps for each point x_i of the image.

The method operates by defining the window (kernel) from geometric (spatial radius) and photometric (color or pixel value). The pixels in the window were neighbors. The window slides in the image and analyzes the neighborhood's mean, then shifts the target pixel with the mean value. These processes will iterate until the mean value is not changed or convergence, which corresponds to the object of interest. The resulting image is smoothed, and the noise is reduced. Mean shift filtering is highly adaptive to object shape and size, which makes it an extensive method for detecting and tracking in computer vision applications (Khattak, Raja, Anjum, and Qasim, 2015). The kernel size only controls the mean-shift and requires less manual intervention than other algorithms (Zhou, Wang, and Schaefer, 2011). Moreover, this method is computationally efficient, fast processing, and robust to variations of object shape (Dorin Comaniciu and Meer, 2002; Xiang, 2009).

While the mean-shift method has several advantages, limitations should be considered. The convergence to the local than the global maximum, especially in images with multiple objects. The performance of the mean-shift method was impacted by the parameter (bandwidth), which is based on the size and contrast of objects (Morales et al., 2011). However, the size of the mean-shift kernel is essential. However, there is no method for selecting the optimum kernel size (Collins, 2003). Finding an optimal kernel size of mean-shift filtering to the interest object is challenging.

2.5 K-means clustering

Clustering is a technique for finding subgroups or clusters in the dataset based on the homogeneous character of the dataset without training or supervision. K-means clustering is one of the most popular unsupervised classifications. K-means clustering is an approach for partitioning data into clusters. The K-means clustering algorithm requires the user to specify the desired number of clusters (*k*). It then assigns each observation in the dataset to one of the *k* clusters. The objective of K-means clustering is to minimize the summation of within-cluster variation.

Gareth, Daniela, Trevor, and Robert (2015) summarized the principle of K-means clustering as follows. The within-cluster variation $W(C_k)$ is estimated from the different observation data and a mean value of the cluster, the calculation starts from:

$$\underset{C_1,\ldots,C_k}{minimize}\left\{\sum_{k=1}^{K} W(C_k)\right\} \tag{5}$$

where C_k contains the indices of the observation data in each cluster. The sum of within-cluster variation could be as small as possible. The square Euclidean distance was the method for finding the minimized value of equation (5), defined as

$$W(C_k) = \frac{1}{|C_k|}\sum_{i,i'\in C_k}\sum_{j=1}^{p}(x_{ij} - x_{i'j})^2, \tag{6}$$

where $|C_k|$ is the number of observation data in cluster k and p is the number of features of observation data. Then replace the Euclidean function with the mean for feature i in the cluster C_k as expressed in (7), and the within-cluster variation function becomes (8)

$$\bar{x}_{kj} = \frac{1}{|C_k|} \sum_{i \in C_k} x_{ij} \tag{7}$$

$$\frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 = 2 \sum_{i \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{kj})^2 \tag{8}$$

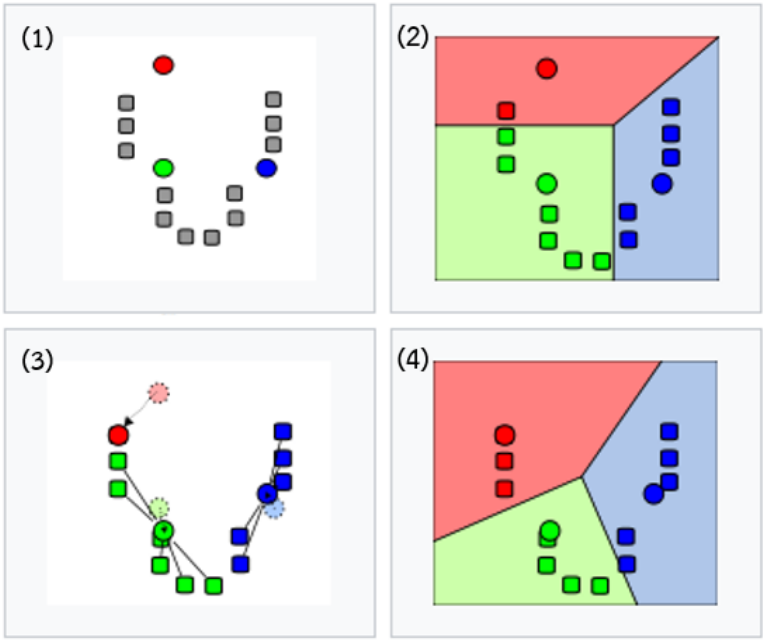
K-means clustering aims to minimize the summation of within-cluster variation in equation (6). The processes for solving the problem are described as follows:

(1) Define the number of clusters (*k*), then random *k* numbers as an initial cluster centroid. Figure 2.6 (1) represents the process of clustering data with *k*=3; red, green, and blue points are the random initial centroids for each cluster.

(2) Defined Voronoi diagram from the partition of the distance between cluster centroids and created boundary around the centroids. Regions of each centroid are presented in color corresponding to the color of the centroid, as shown in Figure 2.6 (2). Assign the observation data to the closest centroid; the data within the same region are clustered.

(3) Compute the new cluster centroid for each cluster. The cluster's centroid is a new mean value of the observation in the cluster of p features. The new centroids moved to a new position, as presented in Figure 2.6 (3).

(4) Repeat processes (2) and (3) until the cluster centroid does not change, resulting in the final clustering process presented in Figure 2.6 (4).



Source: Weston.pace (2007)

Figure 2.6 Processes of K-means clustering.

Even with the process's simplicity and fewer tuning parameters, K-means clustering has some problems. The random initialization trap is the main problem for this method. The random value for the initial centroid is essential to find the optimum centroid, but it can be a local optimum. Figure 2.7 shows an example of the random initialization trap. The three clusters' results were presented in different colors, sizes, and members of each cluster were different due to the initial centroid. The within-cluster variation value was calculated based on the centroid and member of the cluster. These plots in Figure 2.7 showed that the final cluster with the within-cluster variation value of the same dataset processing multiple times produced different clusters with distinct within-cluster variation values. Thus, processing multiple times with different random initial centroids is necessary. Then select the best solution with the smallest within-cluster variation value to solve the problem.

Source: Gareth et al. (2015)

Figure 2.7 K-means clustering of the same dataset with *k=3*. Each plot shows the clustering result using a different set of initial centroids, and the number above the plot is the within-cluster variation.

2.6 Previous studies of weed detection

This section provides a review of weed classification using UAV images, summarizing the studies conducted by other researchers in this field. It highlights the input features, processes, and results employed in these studies. Moreover, the gaps of the previous are discussed.

Peña, Torres-Sánchez, de Castro, Kelly, and López-Granados (2013) developed the process for mapping weed in Early-Season Maize Fields. An automated object-based image analysis (OBIA) procedure was developed using a six-band multispectral camera (visible and near-infrared range) on a series of UAV images to generate a weed map for a maize field. The procedure consists of three phases: 1) classifying crop rows using a dynamic and auto-adaptive classification approach, 2) distinguishing between crops and weeds based on Normalized Difference Vegetation Index (NDVI) threshold and crop rows, and 3) generating a weed infestation map in a grid structure. The estimation of weed coverage through image analysis yielded promising results, with R-squared value of 0.89 and a root mean square error of 0.02. A map categorizing weed coverage into three categories was produced, achieving an overall accuracy of 86%. In the experimental field, 23% of the area was weed-free, and 47% had low weed coverage (<5% weeds), indicating significant potential for reducing herbicide application or other weed-related operations. Nevertheless, the classification of vegetation and background through the use of NDVI number is subject to variation depending on the specific image and geographical area.

The study of Pérez-Ortiz et al. (2015) introduces a weed mapping system in sunflower crop that utilizes vegetation indices derived from RGB and NIR sensors, as well as crop rows, with various classification methods including K-means, Repeated K-means, K-nearest neighbor, linear SVM, kernel SVM, and semi-supervised SVM. The semi-supervised method is developed based on SVM classification and incorporates both labeled and unlabeled data. The results demonstrate that the semi-supervised method achieves the highest performance, and the inclusion of NIR data contributes to improved classification accuracy. However, the semi-supervised classification requires training data, even if the study mentions using a smaller amount of data.

The study of P. Lottes, R. Khanna, J. Pfeifer, R. Siegwart, and C. Stachniss (2017) focuses on developing a classification system for identifying weeds in sugar beet fields. The system incorporates an index to differentiate vegetation from the background and utilizes crop rows in the classification process. The classifier is based on RF

classification. The testing involved three different sensors: two RGB sensors and one RGB+NIR sensor. The system effectively identifies sugar beets and weed species in images of sugar beet farms, benefiting from crop rows and spatial relationships. It performs well even in challenging conditions such as overlapping plants and can detect weeds within the rows. Furthermore, the inclusion of an additional NIR channel improves vegetation detection and classification performance. The study utilizes very high-resolution images, ranging from 0.2 to 5 mm/pixel, captured at low altitudes (2-15 m). These images contain a significant amount of data and require substantial processing time. This method also requires threshold of index for separate background and training data for training RF classification model, which should be done manually.

The study of Gao et al. (2018) focuses on developing a classification process for mapping weeds in early growth state maize fields from RGB images. The method utilizes pixel-based identification of crop rows and detects inter-row weeds, which are automatically labeled and used as training data for RF classification. The segmented image is then inputted into the RF classification to classify intra-row weeds. The process incorporates various geometry features such as length-to-width ratio and asymmetry, obtained through the OBIA technique. The overall accuracy (0.945) and Kappa value (0.912) metrics demonstrate that the RF classifier exhibits strong generalization ability. The OBIA approach was employed for segmentation and feature generation, resulting in high-quality inputs for image classification. However, users are required to manually set parameters, select specific information, and perform the processing steps. The input image has a high resolution of 1.78 mm/pixel, which requires significant time for acquisition and analysis.

De Castro et al. (2018) developed an automatic RF-OBIA algorithm for early weed mapping between and within crop rows using UAV imagery. In this study, an automatic classifier based on RF and OBIA approach was developed and tested in the early stages of sunflower and cotton cultivation. The image underwent segmentation, Digital Surface Model (DSM) extraction, and shadow removal. Vegetation was differentiated from the soil using the NIR/G ratio, which was automatically calculated.

Crop rows were detected by analyzing the length-to-width ratio of the targets after merging. Weeds outside the crop rows were identified based on their location, while weeds within the crop rows were detected using RF classification. The RF model was trained using a predefined dataset of classes, and object features obtained from the segmentation were used for classification. The combination of UAV imagery and RF-OBIA showcased in this study enables accurate weed mapping, both between and within crop rows, with a high weed detection accuracy of 87.9%. The images were captured with a high overlap of 93% to facilitate 3D reconstruction of the crops and obtain crop height information. However, this approach is time-consuming due to the extended duration required for flying the UAV and the subsequent mosaic process. The input features for classification are derived from segmented objects, requiring users to manually set parameters, select specific information, and perform processing steps.

Louargant et al. (2017) evaluated hierarchical self-organising maps for weed mapping using UAV multispectral imagery. This study introduces hierarchical map classifiers for mapping the spatial distribution of *S. marianum* weed. The classifiers utilize features derived from a combination of spectral information from multispectral images and textural information. Three hierarchical map classifiers, namely Supervised Kohonen Network (SKN), Counter-propagation Artificial Neural Network (CP-ANN) and XY-Fusion network (XY-F) were employed to classify the data into *S. marianum* and other plants. The results demonstrate that the CP-ANN classification map achieved the highest accuracy, reaching 98.87%. Nevertheless, the input image resolution of 0.5 m is not suitable for detecting small patches of weeds.

Gašparović et al. (2020) developed an automatic method for weed mapping in oat fields based on UAV imagery. In this study, four classification algorithms were used to create weed maps: manual RF (using pixel-based RF and object-based methods) and automatic RF (using pixel-based and object-based methods). These algorithms applied for weeds and bare soil extraction. The automatic methods used the top 0.1% of NGRDI values for weeds and BI values for bare soil as training data sets. The K-means algorithm was then applied to estimate the presence of weeds and bare soil in areas

without weeds or soil. The automatic object-based classification method achieved the highest accuracy, with an overall accuracy of 89.0% for subset A and 87.1% for subset B. The automatic classification methods were well-developed, using at least 0.25% of the scene size as the training dataset to ensure optimal performance of the random forest classification algorithm in all expected scenarios. The automatic training approach in this study, selecting 0.1% of the index, may be affected by color variation and noise. The presence of white flowers of weed interferes with the BI index, as mentioned by the authors.

Khan et al. (2021) developed a semi-supervised framework for UAV based crop/weed classification. In this study, the development of an optimized semi-supervised learning approach is proposed, offering a method for crop and weed classification at early growth stage. The proposed algorithm based on Generative Adversarial Network (GAN), the framework consists of a generator that provides extra training data for the discriminator, which distinguishes weeds and crops using a small number of image labels. The proposed system was evaluated on the RGB images of pea and strawberry plots. Nonetheless, the inclusion of training data remains essential, even if the study mentions utilizing a smaller dataset. The outcomes of this approach involve sub-images that are categorized as either weed or crop, rather than providing precise pixel coordinates for weed locations.

Su et al. (2022) analyzed spectral and developed process for mapping of blackgrass weed by leveraging machine learning and UAV multispectral imagery. In this study, spectral indices based on multispectral data were generated. The RF classifier with Bayesian hyperparameter optimization was employed to classify blackgrass weed in wheat fields. The results demonstrated high accuracy, especially when using the triangular greenness index (TGI) composed of Green-NIR. The feature selection process reduced the number of features while yielding better results compared to using all the produced features. Additionally, incorporating spatial information from the Guided Filter enhanced the classification outcome by improving results and reducing noise.

However, the classification based on RF requires training data for training classification model.

Previous research has shown that weed mapping in crop fields is crucial for decision-making and management. Researchers are actively exploring new methods for weed detection and trying to find suitable parameters for classification. These efforts indicate a shift towards semi-automatic or fully automatic weed detection processes. However, most weed mapping methods rely on supervised classification, requiring training data. Thresholding methods, on the other hand, depend on variable indices affected by factors like area, time, and lighting conditions, making it difficult to determine optimal values. This research aims to develop a semi-automatic classification process tailored for classifying objects in cassava fields. It combines various preprocessing and classification techniques to establish an efficient workflow.