


## Article

# Rice Growth Stage Classification via RF-Based Machine Learning and Image Processing

Rodney Tai-Chu Sheng <sup>1</sup>, Yu-Hsiang Huang <sup>2</sup>, Pin-Cheng Chan <sup>2</sup>, Showkat Ahmad Bhat <sup>3,\*</sup>, Yi-Chien Wu <sup>4</sup> and Nen-Fu Huang <sup>1,\*</sup> 

<sup>1</sup> Department of Computer Science, National Tsing Hua University, Hsinchu 30013, Taiwan

<sup>2</sup> Institute of Information Systems and Applications, National Tsing Hua University, Hsinchu 30013, Taiwan

<sup>3</sup> College of Electrical Engineering and Computer Science, National Tsing Hua University, Hsinchu 30013, Taiwan

<sup>4</sup> Taichung District Agricultural Research and Extension Station, Changhua County 515008, Taiwan

\* Correspondence: s108064891@m108.nthu.edu.tw (S.A.B.); nfhuang@cs.nthu.edu.tw (N.-F.H.)

**Abstract:** Rice is one of the most significant crops cultivated in Asian countries. In Taiwan, almost half of the arable land is used for growing rice. The life cycle of paddy rice can be divided into several stages: vegetative stage, reproductive stage, and ripening stage. These three main stages can be divided into more detailed stages. However, the transitions between stages are challenging to observe and determine, so experience is required. Thus, rice cultivation is challenging for inexperienced growers, even with the standard of procedure (SOP) provided. Additionally, aging and labor issues have had an impact on agriculture. Furthermore, smart farming has been growing rapidly in recent years and has improved agriculture in many ways. To lower the entry requirements and help novices better understand, we proposed a random forest (RF)-based machine learning (ML) classification model for rice growth stages. The experimental setup installed in the experiment fields consists of an HD smart camera (Speed-dome) to collect the image and video data, along with other internet of things (IoT) devices such as 7-in-1 soil sensors, a weather monitoring station, flow meter, and milometer connected with LoRa base station for numerical data. Then, different image processing techniques such as object detection, object classification, instance segmentation, excess green index (EGI), and modified excess green index (EGI) were used to calculate the paddy height and canopy cover (CC) or green coverage (GC). The proposed ML model uses these values as input. Furthermore, growth-related factors such as height, CC, accumulative temperature, and DAT are used to develop our model. An agronomist has been consulted to label the collected different stages of data. The developed optimal model has achieved an accuracy of 0.98772, and a macro F1-score of 0.98653. Thus, the developed model produces high-performance accuracy and can be employed in real-world scenarios.

**Keywords:** machine learning; image processing; smart farming; paddy rice growth stages; random forest; precision agriculture



**Citation:** Sheng, R.T.-C.; Huang, Y.-H.; Chan, P.-C.; Bhat, S.A.; Wu, Y.-C.; Huang, N.-F. Rice Growth Stage Classification via RF-Based Machine Learning and Image Processing. *Agriculture* **2022**, *12*, 2137. <https://doi.org/10.3390/agriculture12122137>

Academic Editor: Maciej Zaborowicz

Received: 30 October 2022

Accepted: 10 December 2022

Published: 12 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

All human existence is affected by agriculture, which is one of the most significant industries in the world. Besides providing security, nutrition, and health for a nation's citizens, agriculture production is vital for its economy [1]. Paddy rice is one of the most common foods and is cultivated primarily in Asia. In Taiwan, rice is usually cropped from February to November, and is considered the most farmed crop [2]. The first farming period (or cropping season) is from February to July, and the second is from August to November. The second farming period would be shorter than the previous one as both temperature and day length are longer. However, the growth duration of rice plants varies with the variety of species and the environment in which they grow. Based on the season and different species, the life cycle of the rice plant is around 100 to 210 days [3].

The life cycle of rice generally includes the vegetative stage, reproductive stage, and ripening stage. Each stage can be subdivided into more detailed stages. Figure 1 demonstrates the growth stages [4]. In Taiwan, Good Agricultural Practice (TGAP) publishes rice farming methods (or SOP) for farmers [5]; the SOP is more in line with Taiwan's planting conditions. The SOP contains the rice growth stages and related field work according to a particular growth stage (cf. Figures 1–3).

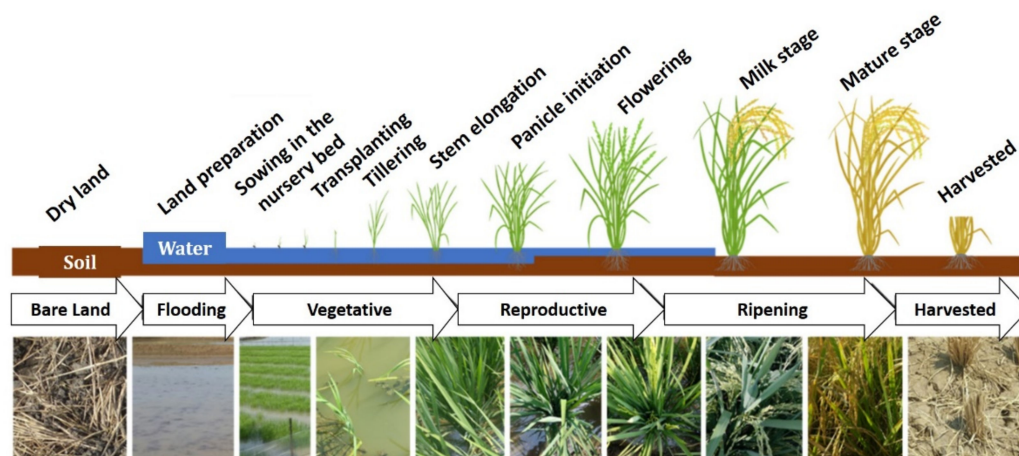


Figure 1. Paddy Rice Growth Stages Illustration from [2].

#### 1<sup>st</sup> cropping season/ farming period

Fertility days	Phase	Stage	Fertilizing	Irrigation	Disease Management
0	Vegetative phase	Seedling stage	1. seedbed preparation 2. basal fertilizer (N:18~24 kg/ha P:50 kg/ha K:20 kg/ha)	flooding irrigation 1 cm	
7		Initial tillering stage	used herbicides	flooding irrigation 3~5 cm	White-tup nematode
14					
28		Active tillering stage	First top dressing (N:36~42 kg/ha)	flooding irrigation 3~5 cm	Rice blast White-tup nematode Rice bakanae disease
35			Second top dressing (N:36~42 kg/ha)		
42	Reproductive phase	Max tillering number stage		sun drying (dry crack in the field)	Sheath blight Rice bakanae disease Rice blast
49				flooding irrigation 3 cm	
56		Panicle formation stage	earring fertilizer (N:18~24 kg/ha K:30 kg/ha)	flooding irrigation 5~10 cm	Rice bakanae disease Sheath blight
63		Booting stage		intermittent irrigation 3 cm	
70		Heading		flooding irrigation 5~10 cm	Rice blast
77	Ripening phase	Milk-ripe stage			
84		Dough-ripe stage		intermittent irrigation 3 cm	
91		Yellow-ripe stage			
98		Maturity stage			
105					
112					
119					

Figure 2. The First Rice Farming Period and Related SOP.

The figures above show some curves or trend lines. From left to right, curved lines represent tillering counts, panicle counts, and panicle weights; straight lines represent paddy weights. There are several common indices for paddy rice, which indicate the growth of paddy rice. The indices, however, cannot be measured with images, they must be measured manually.

The stages after transplantation are sorted into Table 1, which contains 11 stages across three phases. Splitting phases into multiple stages is ideal in farming as different actions are taken at different periods. Based on the consultations with experts, we recommend treating the ripening stage as a single phase. This is because the transitions in ripening phases are extremely difficult to classify precisely. In addition, there are no fields of work to operate in the ripening phase. In this case, the stages are selected across vegetative and reproductive phases, and the ripening phase (treat milk stage, dough stage, yellow ripening

stage, and mature stage as one), in a total of 8 stages (cf. Table 1). These 8 stages are used for the classification task.

### 2<sup>nd</sup> cropping season/ farming period

Fertility days	Phase	Stage	Fertilizing	Irrigation	Disease Management	
0	Vegetative phase	Seedling stage	1. seedbed preparation 2. basal fertilizer ( N:18~24 kg/ha P:50 kg/ha K:20 kg/ha )	flooding irrigation 1 cm		
4						
7		Initial tillering stage	used herbicides	flooding irrigation 3~5 cm	White-tup nematode	
14					White-tup nematode	
21		Active tillering stage	First top dressing (N:39~45 kg/ha)	flooding irrigation 3~5 cm	Sheath blight Rice bacterial blight	
28			Second top dressing (N: 39~45 kg/ha)			
35		Max tillering number stage		sun drying (dry crack in the field)	Sheath blight Rice bacterial blight	
39	Reproductive phase	Panicle formation stage		flooding irrigation 3 cm		
42			earring fertilizer (N:19~26 kg/ha K:36 kg/ha)	flooding irrigation 5~10 cm		
49		Booting stage		intermittent irrigation 3 cm		
70	Ripening phase	Heading		flooding irrigation 5~10 cm		
77		Milk-ripe stage				
84		Dough-ripe stage		intermittent irrigation 3 cm		
91		Yellow-ripe stage				
98		Maturity stage				
105						

**Figure 3.** The Second Rice Farming Period and Related SOP.

**Table 1.** Growth Stages of Paddy Rice Sorted by TGAP and Our Experts for Classification Task.

Stage	Phase	Selection
Transplantation	Vegetative	Yes
Early Tillering		Yes
Peak Tillering		Yes
Late Tillering		Yes
Panicle Initiation	Reproductive	Yes
Booting		Yes
Heading		Yes
Milk	Ripening	Yes
Dough		Yes
Yellow Ripening		Yes
Mature		Yes

Although the farming SOP covers most of the cultivation, the transition between stages is still difficult to observe. Typically, farmers determine stages by observing changes in leaf color, fertility days, and physiological changes. These methods are effective for experienced farmers, but not every farmer has had the same experience. However, it is difficult to determine the transitions between each stage. Thus, the cultivation of rice is challenging and experience-intensive, leading to labor issues (aging) and a lack of experience inheritance. If an inexperienced person only follows the provided SOP, they are likely to fail. Different field studies need to be carried out at different stages, and adjustments need to be made according to the actual situation. Although it is difficult, replacing the younger generations is more challenging.

Introducing smart farming [6,7] has led to an improvement in yield production [8] and cropping period. In addition, deep learning or machine learning [9] is increasingly being used in smart farming, such as for pest detection, yield prediction, and crop quality. In this paper, we propose a classification model to classify the paddy rice growth stages. As a result, farmers and young generations will have a better understanding of their crops, improve their farming methods and, most importantly, lower entry requirements. Thus, agricultural operations can be monitored and controlled remotely using the developed system rather than by observing them onsite and observing how they are performing. Through

smartphones or desktops, farmers can stay up to date with paddy rice developments. As a result of structural and environmental changes such as aging farmers, low birthrates, worker shortages, difficulty in farm maintenance, and experience inheritance disconnection in recent years, smart farming is more relevant to Taiwan's agriculture industry. The developed system can make farming easier, as the majority of fieldwork can be automated, reducing the amount of manual work. As it does not require patrolling, with 24/7 monitoring changes are more intuitive. It also provides a beginners' guide to understanding crops and making informed decisions, allows workers to be trained more efficiently and lowers labor costs and is also an eco-friendly approach to sustainable agriculture or permaculture.

The objective of this study is to explore the possibilities of using IoT and big data (BD)-driven ML prediction models to assist farmers in efficiently determining different growth stages of paddy rice and in determining grain yield. It is very important to explore alternatives to traditional agriculture systems in view of the continuously increasing food demand of the world. In addition, accurate estimation of the heading date of paddy rice greatly helps farmers to understand the adaptability of different crop varieties in a given location. For the calculation of the paddy height and CC or GC, different image processing techniques are used: analytics for the calculation and classification of different growth parameters using machine learning in computer vision; image recognition for individual plants; automated plant height calculation; automated green coverage calculation is employed to monitor the paddy's tillering status; and automated leaf color classification for monitoring the paddy's nitrogen level.

## 2. Background and Related Work

Smart Farming and IoT are expected to have a significant impact [10], especially on rice production, through BD [9], ML, and the IoT [11]. Smart farming for rice strategies can now anticipate changes and spot opportunities [12]. This is due to the increasing volume and diversity of data being collected and acquired by these advancing IoT technologies. A high-quality sensor dataset has a significant impact on the effectiveness of the modeling procedures using deep learning (DL) algorithms. Due to the extensive use of these three elements (for example, BD, ML, and IoT) in rice production operations in agriculture, a new era of rice smart farming or rice precision agriculture has emerged. There have been several studies in the literature that study the use of these technologies in paddy cultivation. Researchers in the study in [13] used image processing and machine learning algorithms to estimate or predict the maturity of rice panicles to determine the appropriate harvest time. The authors collected paddy panicles from two rice varieties and scanned them with a flatbed scanner after dehulling. Then, they segmented the paddy panicle based on the hue, saturation, and intensity (HSI) color space. After that, they used RGB and HSI color spaces to extract color features. A total of 22 color features were extracted from each paddy panicle image. They developed an RF model to predict maturity. The model was trained with the 22 color features and 10-fold cross-validation was applied. They used R-square and root mean square error (RMSE) to evaluate their model.

In comparison to the variation in colors and standard deviation, the correlation between the mean value of color and maturity is relatively low. Then, they used these optimal features to repeat the training process. Finally, an R-squared of 0.93, RMSE of 1.18%, R-squared of 0.94, and RMSE of 1.60% were obtained on the two varieties they collected, respectively. The authors claimed that their methods performed better than other research [14,15]. However, the proposed method was only validated for the two rice varieties they had experimented with.

Studies [16,17] aimed to estimate rice growth, nitrogen (N) nutrition status, and canopy cover. For extracting color indices, they analyzed digital camera images and image processing methods. CC is commonly used to determine crop yield and growth parameters, such as the leaf area index (LAI). The process of measuring CC on the field is ineffective and takes a considerable amount of time. To distinguish the plant pixels from the background pixels, the authors used color indices. They calculated CC values as the ratio of the plant



pixels to the total pixels. Among the color indices, the excess green index (EGI) and modified EGI (MEGI) are effective for measuring rice CC in the field. The authors adopted the CC method and developed step-wise multiple linear regression (SMLR) to estimate crop growth and N nutrition. They estimated the LAI, the dry weight, and the nitrogen accumulation. According to their study, CC has positive correlations with LAI, shoot dry weight, and shoot N accumulation. Their models for LAI, shoot dry weight, and shoot N accumulation showed acceptable precision and accuracy. The authors state that digital cameras can be used to monitor rice growth non-destructively.

In another study [18], SVM was used to detect the age of rice farms. The RGB variables were obtained using histograms generated from drone-captured images. Rice growth can be divided into four stages, including the initial phase, the vegetative phase, the generative phase, and the harvest phase. Based on their results, 93.3% accuracy was achieved. By using drone images, they claimed to be able to identify the age-growth stage of rice quickly using histograms and SVM. A time span ranging from weeks to months was defined by the authors for the stages. Their classification result is acceptable given that they only considered five stages (post-harvest included).

Another study [19] proposed an automatic identification algorithm to determine whether the rice has entered the tiller period. They had installed some surveillance cameras for data acquisition. The proposed rice tiller recognition method combines principal component analysis (PCA) and Support Vector Machine (SVM). The plants are first segmented from the background using Excess Green Feature and then go through K-means clustering for further segmentation. Then, more features are extracted and data dimensions are reduced using the PCA algorithm before undergoing the SVM binary classification model. These results are claimed to be promising and generalizable.

Satellite images are often used to determine vegetation cover or canopy cover. Another research [20] developed a paddy growth stages classification by adopting LANDSAT-8 spectral data, and deep learning methods with multiple regularizations (Fast Dropout, Dropout, and Batch Normalization). The vegetation indices used are the Enhanced Vegetation Index (EVI), the Normalized Difference Vegetation Index (NDVI), the Atmospherically Resistant Vegetation Index (ARVI), and the Land Surface Water Index (LSWI).

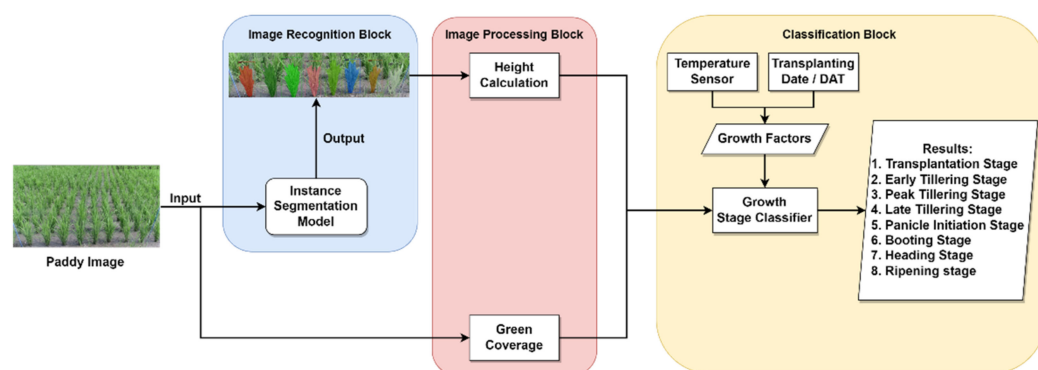
Aging and the declining number of farmers are impacting Japan's agriculture [21], therefore the aim of this study was to support farmers by estimating the paddy's growth stage. The authors had their images taken at different heights such as 30 m, 60 m, and 100 m. They attached a multi-spectral sensor to their drones to capture pictures of multiple wavelength bands. The adopted multi-spectral sensor is capable of capturing Green, Red, Red Edge, and Near Infrared images. To generate NDVI images, they used the reflectance of Red and Near Infrared images. The authors implemented a Convolutional Neural Network (CNN) to classify five paddy growth stages. They concluded that images taken at a height of 60 m have the most accurate results.

Image processing techniques were mostly used in the research discussed above to obtain color features or indices of observation objects, and ML algorithms to identify growth stages or maturity levels. The methods they use are convincing and have proven effective. However, to access or collect the data, human labor is required, which is time-consuming and less automated. Moreover, many factors could affect or reflect rice growth, such as temperature, height, fertility days, or days after transplanting (DAT). Furthermore, the time intervals of [18] are large, which makes it unlikely to determine the precise growth stages successfully. This study makes the following contributions: (1) developed a classification model to classify the rice growth stage, using an RF-based ML model. (2) an height calculation method using OpenVAS camera calibration method was developed to calculate paddy rice height, (3) image processing techniques such as excess green index (EGI) and modified excess green index (MEGI) techniques and instance segmentation were used to determine canopy cover, paddy rice maturity, and the nitrogen level of the paddy rice plants, which were then given as inputs to the developed classification model, and (4) agricultural operations can be controlled and monitored remotely rather than by observing them in

person through the developed agriculture platform. The results show that the transition between paddy rice growth stages can be determined and classified using ML techniques. Therefore, the proposed model has practical significance and is innovative.

### 3. Design and Implementation

For paddy rice, there are three common methods to determine the growth stage: (a) days after transplantation (DAT), (b) the age of leaves, and (c) the physiological observation. In comparison to counting the age of leaves, DAT and physiological observation are more convenient and faster. Counting the age of leaves provides the most accurate estimation, but it is time-consuming and expensive. Our classification system model was based on DAT and exterior observation (images). Figure 4 illustrates the system architecture including image recognition, image processing, and classification, which will be discussed in this section. To extract paddy instances from an image, a segmentation model is developed. The paddy height can be calculated using the output results of the model and camera calibration technique. Excess Green Index (ExG/EGI) is a common method of describing crop CC and has shown high performance in experiments. The canopy cover rate was calculated using the EGI method. In addition to other features, the height calculation and canopy coverage rate are inputs to the classification model. By analyzing the inputs, the model will eventually determine the growth stage. An expert's advice is sought when selecting criteria for growth stages.



**Figure 4.** The system Architecture for the Paddy Rice Growth Stage Classification.

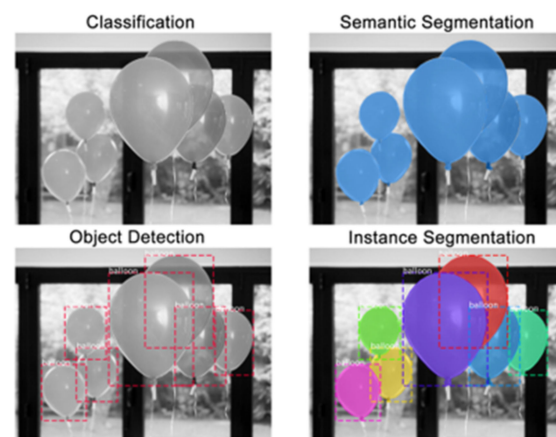
#### 3.1. Image Recognition

Image recognition tasks include object detection, object classification, semantic segmentation, instance segmentation, and panoptic segmentation. Depending on the user's needs, each task serves a different purpose. In this paper, instance segmentation is used because it allows the user to calculate the height, analyze leaf color, and so on for every predicted instance.

#### 3.2. Instance Segmentation Model

Instance Segmentation is the technique to segment every detected-and-classified individual object in an image. It identifies object outlines at the pixel level. In general, it detects instances of objects and marks their boundaries.

Detectron2 [22] is a well-known library for detection and segmentation algorithms, which was developed by Facebook AI Research (FAIR). It provides numerous pre-trained models for different computer vision tasks as shown in Figure 5. For this study, the model in the category COCO Instance Segmentation Baselines model with Mask R-CNN was selected. Mask R-CNN with ResNeXt-101-FPN(X101-FPN) [23] backbone has the highest AP among the categories (cf. Table 2). Mask R-CNN [24] is widely used to perform instant segmentation tasks and other tasks as well.



**Figure 5.** Difference Between Computer Vision Tasks (Adopted from [25]).

**Table 2.** Detectron2 COCO Instance Segmentation Baselines Models with Mask R-CNN Scores.

Name	lr Sched	Train Time (s/iter)	Inference Time (s/im)	Train mem (GB)	Box AP	Mask AP	Model Id
R50-C4	1×	0.584	0.110	5.2	36.8	32.2	137259246
R50-DC5	1×	0.471	0.076	6.5	38.3	34.2	137260150
R50-FPN	1×	0.261	0.043	3.4	38.6	35.2	137260431
R50-C4	3×	0.575	0.111	5.2	39.8	34.4	137849525
R50-DC5	3×	0.470	0.076	6.5	40.0	35.9	137849551
R50-FPN	3×	0.261	0.043	3.4	41.0	37.2	137849600
R101-C4	3×	0.652	0.145	6.3	42.6	36.7	138363239
R101-DC5	3×	0.545	0.092	7.6	41.9	37.3	138363294
R101-FPN	3×	0.340	0.056	4.6	42.9	38.6	138205316
X101-FPN	3×	0.690	0.103	7.2	44.3	39.5	139653917

### 3.3. Data Collection and Training Setup

In Taiwan, there are two major paddy-growing seasons: the first farming period (March to July), and the second farming period (August to November). Most farmers had their paddy seedlings transplanted during these periods. Growing species affect the length of the farming cycle. Fast-breeding species can be harvested in 90 to 110 days, and common species can be harvested in 100 to 140 days. The length of the farming period is also affected by farming strategies. There will be a shorter length in the northern Taiwan area, while it will be longer in the central and south Taiwan areas. This is because the weather varies according to geographical location.

To implement and test the developed smart agriculture platform for paddy rice, we collaborated with three major different farms across Taiwan. One field is in north Taiwan (Hsinchu County), and the other two are in central Taiwan (Changhua County). The smart-agriculture platform developed can monitor any devices installed on the farm, including IoT devices (soil sensors, weather stations, etc.) and surveillance cameras. Real-time video streams and sensor data can be viewed on the platform [26]. Surveillance systems based on IoT devices (Speed-dome camera) have been installed in these experimental fields. Images are automatically captured by the camera on a schedule. With a wide temperature range of  $-40$  to  $60$  °C, the Speed-dome can capture clear images at night and has a 360-degree lens rotation and  $33\times$  zoom capabilities. These schedules (including the camera's angle) were already set up in the developed platform [26], and on-site previewing and streaming are also available. The surveillance camera (Speed-Dome) is powered by a solar panel and has

a Raspberry Pi to perform actions (downlink/execute commands from the website, capture images, upload images, record video, etc.). AWS cloud servers control the smart agriculture platform devices and actuators, which record 100 presets of fixed camera angles, perform regular shooting tasks, upload images automatically to Amazon S3 cloud, and establish an image database of farms and crops. Real-time video streams and sensor data can be viewed on the platform. Using both farming periods, we collected 29,324 amounts of data (based on every paddy instance predicted by the instance segmentation model). As we installed equipment on Field 1 much earlier than on the other fields, a large amount of data has been collected from it. Early stages (especially the transplanting stage and early tillering stage) have shorter periods than others, so their data records are smaller, resulting in an imbalanced dataset. Table 3 represents the data collected from the three experimental fields for two paddy growing seasons and the data count for each stage.

**Table 3.** Total data collected from three experimental fields and data count for each stage.

Field	Farming Period	Data Count	Stages	Data Count	Stages	Data Count
1	1, 2	25,140	1	36	5	3865
2	1, 2	600	2	200	6	4592
3	1	3584	3	2316	7	5406
			4	3638	8	9271
Total				29,324		

To increase the data size and reduce overfitting when training machine-learning models, Dectron2-based data augmentation has been used in this study. There are two basic augmentations provided by default in Detectron2: random flipping and resizing the shortest edge. For the training setup, the COCO-Instance Segmentation task was used as the input format. Image per batch is set to 4, the learning rate is 0.01, weight decay is 0.0001, and epoch is set to 30. The training process was carried out on a Google Colab Pro environment which consists of 16 GB of GPU memory.

### 3.4. Image Processing

A description of each algorithm or method used in image processing is presented in this section. Two parts make up this section: a calculation of height as well as a calculation of green coverage.

#### 3.4.1. Height Calculation Using Camera Calibration

The image recognition block's output results include not only segmentation but also the bounding boxes that indicate each of the predicted instances. These coordinates could be incorporated into the following height calculation. A height calculation method using the OpenVAS Camera Calibration method was developed in our lab and has been applied in different scenarios with good results. As a result, this method was applied to determine the height of the paddy plant in this paper.

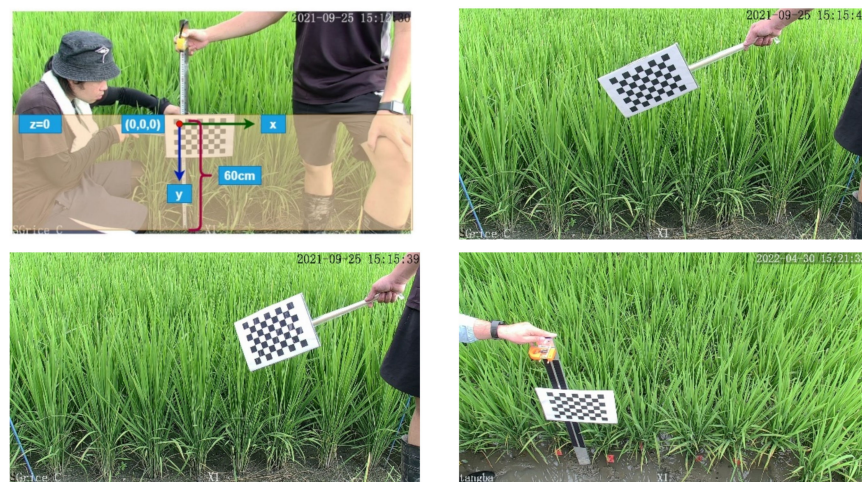
The calibration of a camera is the process of estimating its parameters and determining the relationship between its 3D coordinates in the real world and its equivalent 2D projection in the image captured by the calibrated camera. OpenCV provides a camera calibration method, which simplifies our work greatly. It is based on Zhang's. From a technical standpoint, camera calibration determines the camera matrix, distortion coefficient, rotation vector, and translation vector. The camera matrix transforms 3D world coordinates into 2D image coordinates. The distortion coefficient returns the position of the camera in the world, with the values of the rotation and translation vectors. OpenCV Camera Calibration utilizes a chessboard (whose size is known) to obtain real-world coordinates for 3D points. The calibration steps are as follows:

- Define real-world coordinates of 3D points using a known-size chessboard.



- Capture different viewpoints of the chessboard.
- Find the pixel coordinates ( $u, v$ ) for each 3D point in different images (use findChessboardCorners() method from OpenCV)
- Find camera parameters (use the calibrateCamera() method from OpenCV).

Figure 6 shows the process of defining real-world coordinates. The X and Y axes point to the chessboard, and the Z axis is perpendicular to the chessboard (this can be treated as the distance from the camera to the chessboard, or plane). The red dot represents the known 3D origin point (0, 0, 0). By calculating the camera parameters from known 3D world coordinates ( $X_w, Y_w, Z_w$ ), we can obtain their corresponding pixel location ( $U, V$ ) in the image. By calibrating the camera with the bounding box coordinates ( $U_j, V_j$ ) provided by the Instance Segmentation model shown in Figure 7, their world coordinates ( $X_j, Y_j, Z_j$ ) can be retrieved as well. Lastly, assuming  $Z = 0$ , by calculating the difference between  $Y_w$  and  $Y_j$ , the height data are obtained.



**Figure 6.** Demonstration of Camera Calibration Using a Chessboard and Different Viewpoints Captured.



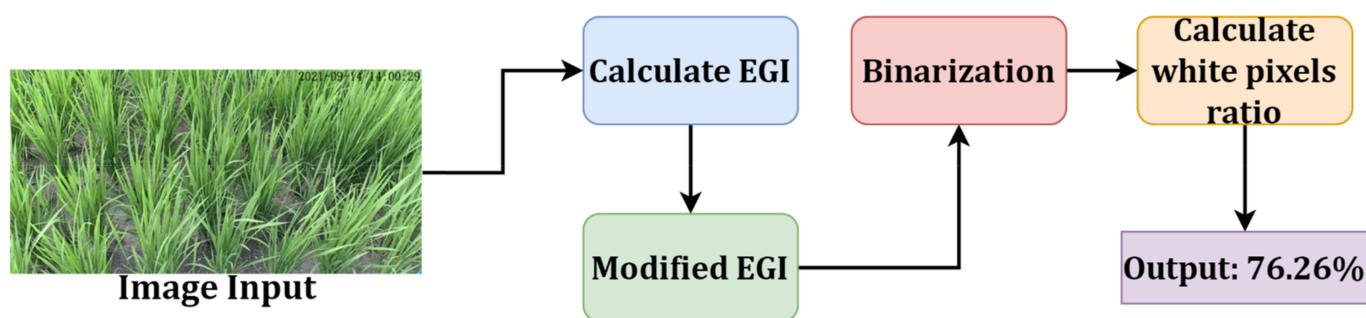
**Figure 7.** The Coordinates of Paddy Instances Predicted by Instance Segmentation Model.

### 3.4.2. Green Coverage (GC) Rate

As mentioned, crop canopy cover (CC) is one of the common ways to represent vegetation. a widely used method of evaluating the rate of CC is EGI, which is a widely used image processing method. The objective of this study was to evaluate green coverage using the EGI methods using CC's concept. The designed procedure is shown in Figure 8.

EGI and MEGI are employed to separate plants from backgrounds. After that, image binarization is carried out to filter some noises to achieve cleaner segmentation. Eventually, the GC rate will be calculated through the white pixels' ratio based on the binarization results. To determine the GC rate, we used the following image selection principles:

- The image is based on a large area of rice field and tries to avoid the non-field area.
- The image angle is based on the depression angle and tries to cover only the rice field area.
- Likely, the uneven coloration or partial brightness of the objects in the image will be caused by the intensity of the sunlight in the morning (sunrise) and the evening (sunset). To reduce the impact of the natural environment on the photos, we chose the images taken between 8:00 am and 4:00 pm.



**Figure 8.** Flow Chart of Calculating GC Rate.

#### 3.4.3. Excess Green Index (EGI) and Modified EGI

Most of the images have three: Red (R), Green (G), and Blue (B), basic channels. The EGI method can be used to filter out or segment plants from the background. The color index of RGB needs to be normalized before the EGI function can be applied. Modified EGI (MEGI) is a variation of EGI. Plant pixels have greater G and g thus greater EGI and MEGI, indicating better segmentation between the paddy rice plant and the background. Since the natural environment is complex, the output results of EGI might not be ideal. In this case, MEGI is adopted for further processing to eliminate non-background pixels.

#### 3.4.4. Binarization and CC Rate Calculation

Image binarization is a common image processing technique that transforms the image into black and white or 0 and 255 in values. MEGI outputs are not necessarily optimal for most images. In that case, the role of binarization here is to remove smaller MEGI values which potentially might be background pixels. However, it also simplifies the calculation of the ratio of white pixels as well. The CC rate is calculated based on the white pixel ratio against the image's size as follows:

$$\text{GC Rate} = (\text{WhitePixels}) / (\text{Width} \times \text{Height}) \times 100\% \quad (1)$$

where width and height indicate the image's width and height.

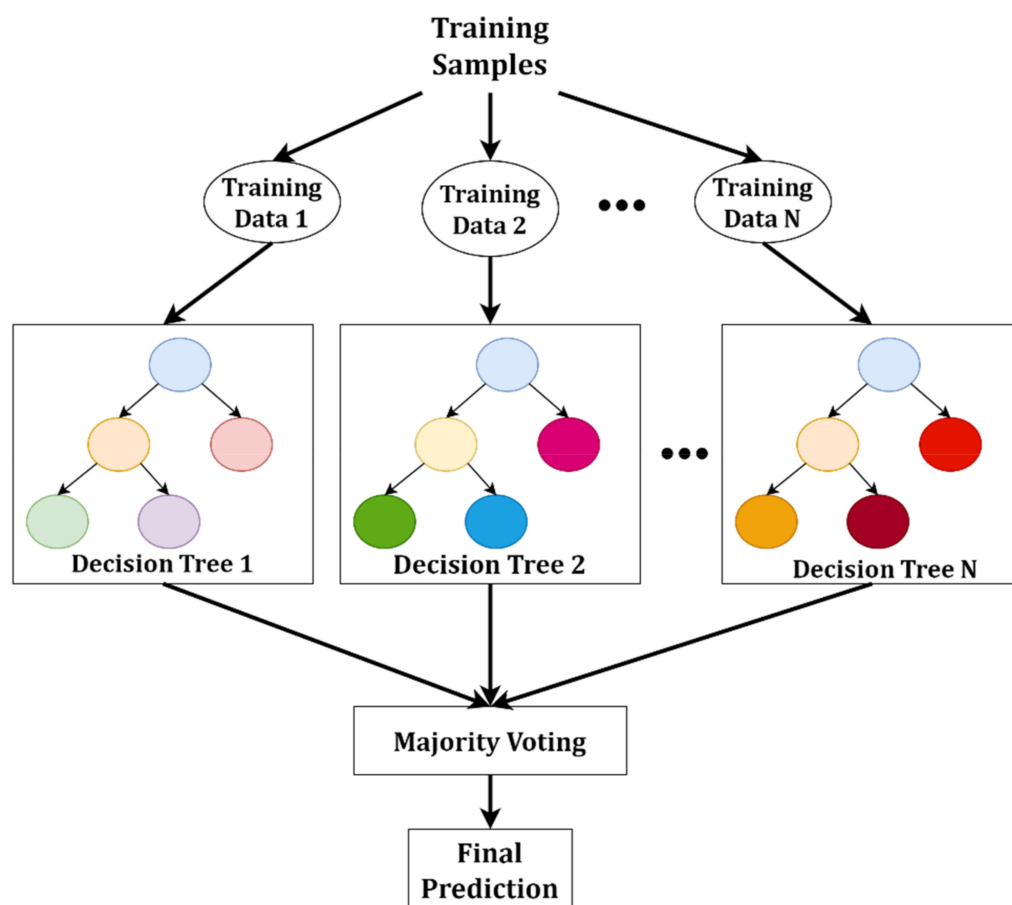
#### 3.5. RF-Based ML Classification Model

For classifying the growth stage, a Random Forest (RF) [27] algorithm-based ML model was developed [28]. As an ensemble algorithm, bagging (or bootstrap aggregating) fits multiple models on different subsets and then combines their results. Random Forest is an ensemble of decision trees and is trained with the bagging method. It can be used for classification and regression. Multi-decision trees are built and merged using RF, resulting in more accurate and stable predictions.

RF uses randomly selected samples from the training dataset to train each decision tree, so every tree will be trained with a different subset of samples. As a result, each decision tree becomes more unique and the correlation between trees is reduced. RF aggregates the

predictions from the individual decision trees and takes the majority votes (classification) or average (regression). The decision tree uses the Information Gain or Gini Impurity criterion to split nodes (generate child nodes). Both methods are based on the concept of entropy. Entropy can be defined as a metric of the purity of the sub-split or how much variance the data have (how data are distributed). The decision tree calculates the entropy of each feature after every split, selects the most significant feature, and starts splitting according to it. High entropy means high impurity or many different classes in the branch.

Information gain is calculated for a split by subtracting the weighted entropies of each child node from the root node. Higher information gain means the child node has less entropy and less variance, which is optimal for splitting. Gini impurity is the probability that the data point is incorrectly classified. Consequently, a node with zero Gini Impurity contains only one class. However, these two methods have no significant difference in the performance of the model [29]. In SciKit-Learn [30], the default criterion for RF is Gini Impurity. The output results from the image processing block (height and CC Rate) will be used as inputs for the RF model. The RF regression model (cf. Figure 9) has been selected for classifying the growth stages of paddy rice because it performs well on large datasets and produces more accurate predictions by combining multiple decision trees. In addition, it has a low variance, which is ideal for our dataset since it consists of different species and experimental fields.



**Figure 9.** The architecture of Random Forest.

Each input image might have multiple predicted instances and a certain GC rate. Each instance is considered a record of data, so an image may consist of more than one record due to the instance segmentation model's results. The growth factors are collected from sensors installed in the field. Experts in the field gather ground truth about the field (height, tillering amount, etc.). The following 12 features are included in the training of our model as shown in Table 4.

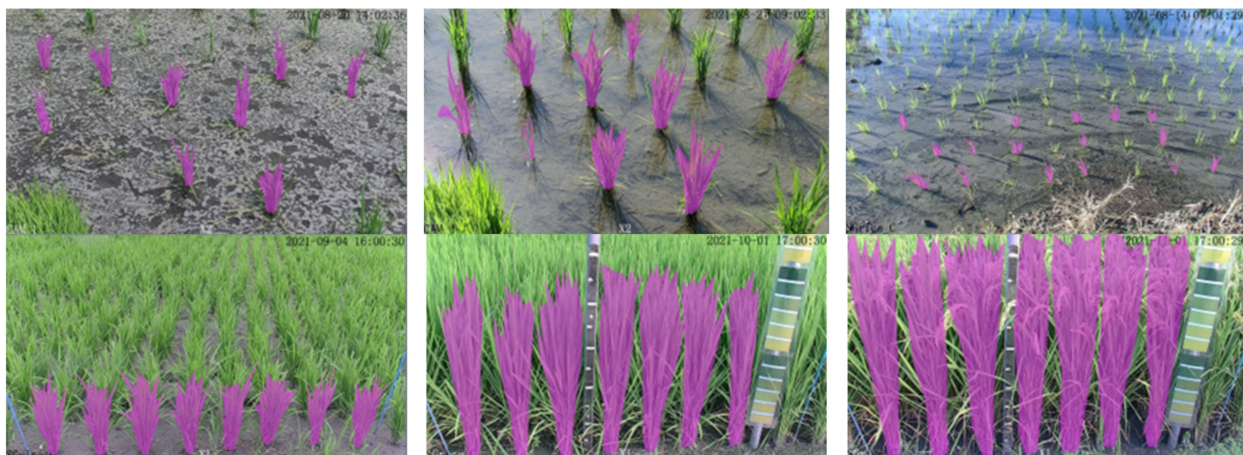


**Table 4.** The Definition of Each Input Training Feature.

Training Features	Definition
green_cov	GC rate calculated from the image, outputted from GC block
period	Farming period (first or second), 1 and 2 are used to represent these growing periods
FirstDay_day	The day of transplantation date (e.g., 2022/03/04, day = 4)
FirstDay_month	The month of transplantation date (e.g., 2022/03/19, month = 3)
FirstDay_year	The year of transplantation data ((e.g., 2022/03/04, year = 2022))
ObsTime_day	The day of the input image's observation date (e.g., 2022/03/19, day = 4)
ObsTime_month	The month of input image's observation month (e.g., 2022/03/19, month = 3)
ObsTime_year	The year of input image's observation year (e.g., 2022/03/19, year = 2022)
DAT	Days after transplantation (Observation Date-First Day (Date) = DAT)
Temperature (°C)	Air temperature of the day collected by installed sensors
acc_temp	Accumulated temperature, T-10 = Effective accumulated temperature of the day, Start summing up from the transplantation day till the observation day
height	Output (height) from height calculation block

### 3.6. Data Labelling

The selection of stages and the labeling of the image dataset to represent different growth stages of the paddy rice from transplanting to maturity were supervised by experts. The following Table 5 represents the selected stages and their codes. A total of 4200 labeled instances were obtained from 607 images (577 were for training, 30 were for testing). For labeling purposes, the leaves on both sides of the main body are not included, thus keeping the body as a cylinder as much as possible. During the process of labeling, we encountered some difficulties. The later stages were the most difficult to observe from the highest point, except for the early stages. There may be some differences in height between the real subjects and the models. In the later stages of tillering, the leaves and panicles sagged and overlapped, making labeling harder. Figure 10 represents some of the labeled image samples.

**Figure 10.** Demonstration of Labelled Images.



**Table 5.** Label Code for Each Growth Stage.

Growth Stage	Code	Growth Stage	Code
Transplantation Stage	1	Panicle Initiation Stage	5
Early Tillering Stage	2	Booting Stage	6
Peak Tillering Stage	3	Heading Stage	7
Late Tillering Stage	4	Ripening Stage	8

## 4. Experiment and Results

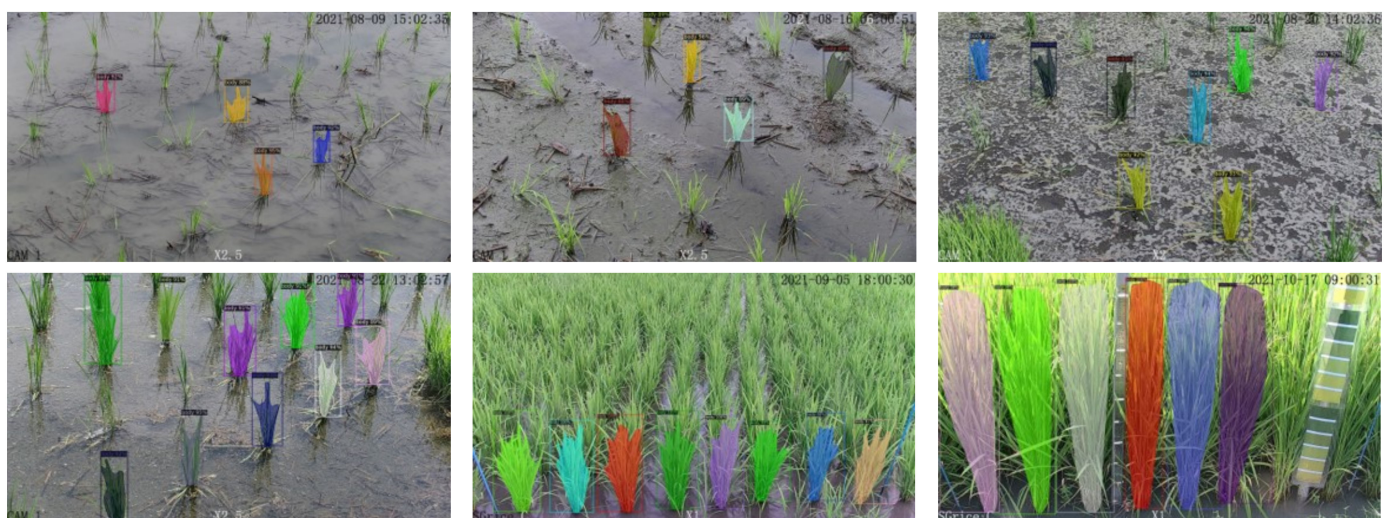
### 4.1. Image Recognition Using Instance Segmentation Model

To evaluate the performance of the model, Average Precision (AP) and Average Recall (AR) are used, both of which range from 0 to 1. In AP tests, there are three common AP scores: AP, AP50, and AP75, which encompass IoU values of 0.5 to 0.95, and 0.5 to 0.75, respectively. There are two types of evaluation in the instant segmentation model: bounding box and segmentation. Both have AP and AR performance scores. The bounding box AP of the developed model is 0.559, and the segmentation AP is 0.506 shown in Table 6. Both of the scores are better than the Detectron2 baseline model. A recall is a measure of how many of the positive cases the classifier correctly predicted over all the positive cases in the data. Because every instance is not labeled, predicting as accurately as possible is the priority. Both AR scores indicate that the model is capable of detecting most instances in the input image. Given the amount of data, the proposed model has an AP and AR that is comparable to or even better than the baseline model from Detectron2.

**Table 6.** Evaluation of Instance Segmentation Model.

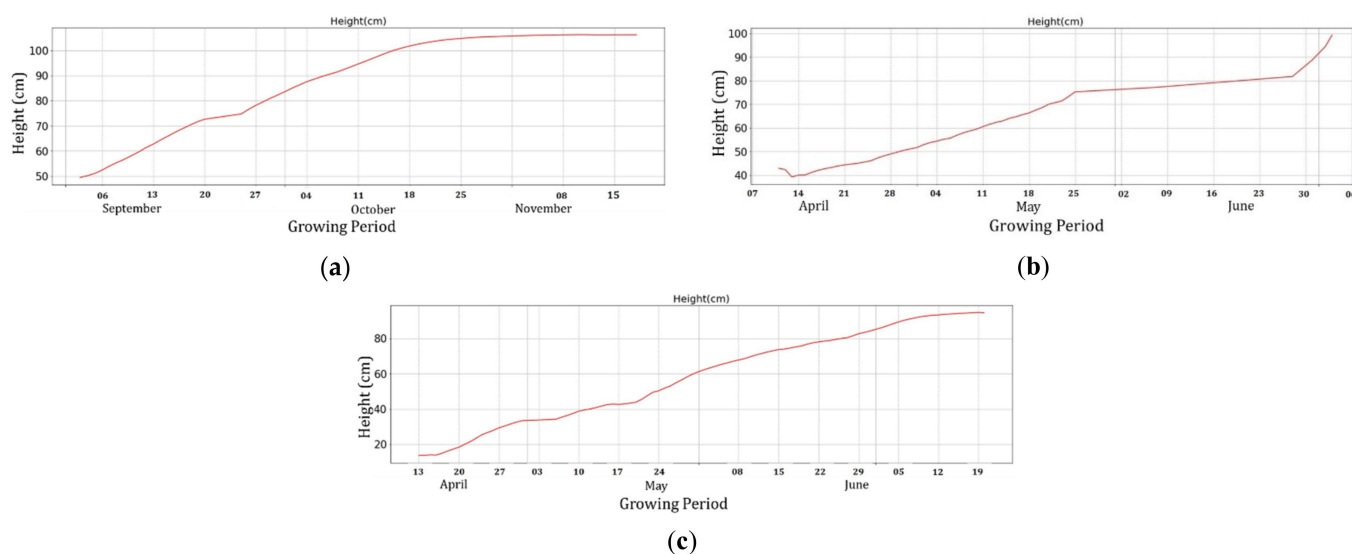
Evaluation Type	Score			
	AP	AP50	AP75	AR
Bounding Box	0.559	0.780	0.671	0.613
Segmentation	0.506	0.780	0.660	0.539

Some predicted results are shown in Figure 11 below. It can be observed that our model can segment paddy instances well. The first two images in Figure 11 are in an early stage because we did not label every paddy instance in the image. Therefore, the number of predicted instances is relatively low.

**Figure 11.** Model's Predicted Instances.

#### 4.2. Image Processing-Based Height Calculation

The height of the paddy instances predicted by the RF-based DL model was calculated using the image processing techniques discussed in the previous section. The height calculation results are shown in Figure 12 throughout the growing seasons. The results are averaged based on the day and smoothed using EWM (com = 5). Note that the trend graph only serves the purpose of monitoring the height trend. The height of each paddy instance is still considered an input record for the classification model. As illustrated in Figure 12c, the results for Field 3 have several gaps in May and June due to bent and drooping leaves and a relatively high shooting angle. This makes it difficult for our model to identify the base of the rice. In addition, weeds blocked the view of the base.



**Figure 12.** The Height Trend Graph of three experimental Fields. (a) Height Trend of Field 1. (b) Height Trend of Field 2. (c) Height Trend of Field 3.

The obtained results were compared with ground truth that was collected by experts, which is shown in Tables 7–9. As the natural curvature of the rice body increased, the error rate increased as well. Additionally, the ground truths were measured manually by pulling and straightening the longest leaf.

**Table 7.** Field 1—Comparison of Height Calculation Based on Image and Ground Truth.

Date.	9/3	9/10	9/17	9/24	10/1	10/8	10/15	Average
Height Calculation	47.8	62.7	75.5	84.7	87.6	82.3	97.1	
Ground Truth	48.0	60.4	75.4	85.9	90.5	97.8	106.0	
Difference	−0.2	+2.3	+1.2	−1.2	−2.9	−5.5	−8.9	3.0
Error Rate	−0.42(%)	+3.81(%)	+0.13(%)	−1.40(%)	−3.20(%)	−5.62(%)	−8.40(%)	3.28(%)

**Table 8.** Field 2—Comparison of Height Calculation Based on Image and Ground Truth.

Date	3/17	3/25	4/1	4/8	4/15	4/22	4/29	5/6	5/13	5/20	5/27	6/3	6/10	6/17	Average
Height Calculation Result	17.38	32.32	37.58	41.30	46.56	57.84	67.26	69.99	76.51	81.57	82.32	92.33	94.95	95.01	
Ground Truth	18.38	33.88	41.38	55.88	56.38	58.75	67.88	69.75	71.75	78.00	81.38	83.38	94.00	103.38	
Difference	−1.00	−1.59	−3.79	−14.58	−9.82	−0.91	−0.62	0.24	4.75	3.57	0.84	8.96	0.95	−8.33	−1.52
Error Rate	5.42%	4.70%	9.17%	26.09%	17.42%	1.55%	0.91%	0.34%	6.63%	4.58%	1.16%	10.74%	1.01%	8.06%	6.98%

**Table 9.** Field 3—Comparison of Height Calculation Based on Image and Ground Truth.

Date	3/12	3/19	3/26	4/1	4/9	4/16	4/23	5/7	5/28	6/3	Average
Height Calculation Result	41.88	45.35	52.76	57.52	63.00	68.04	78.53	86.39	105.03	123.26	
Ground Truth	10.3	18	26.85	33.6	38.1	48.9	58.8	71.2	100.8	103.9	
Difference	31.58	27.35	25.91	23.92	24.90	19.14	19.73	15.19	4.23	19.36	21.13
Error Rate	306.60%	151.94%	96.50%	71.19%	65.35%	39.14%	33.55%	21.33%	4.19%	18.63%	80.84%

As can be observed from Table 9, Field 3 has the largest difference and also has the greatest error rate. The camera height may have affected the calibration since Field 3 has the highest camera height as represented in Table 10. By comparing Field 1 to Fields 2 and 3, it can be observed that there is a noticeable difference. A higher camera height in Field 1 results in more horizontal views than in Fields 2 and 3. As a result of the relationship between camera height and depression angle, the images are less than ideal, and their distortion causes a substantial calculation error rate. Furthermore, the image distortion makes it difficult to visualize the entire paddy instances, resulting in an inadequate inference of the instance segmentation model.

**Table 10.** Camera Calibration Parameters of Three Fields.

Field	Depression Angle	Distance from Camera to Chessboard (m)	Chess Height above Ground (m)	Camera Height (m)
1	23	1.2	0.6	1.0689
2	35	1.45	0.6	1.4317
3	50	1.596	0.6	1.8228

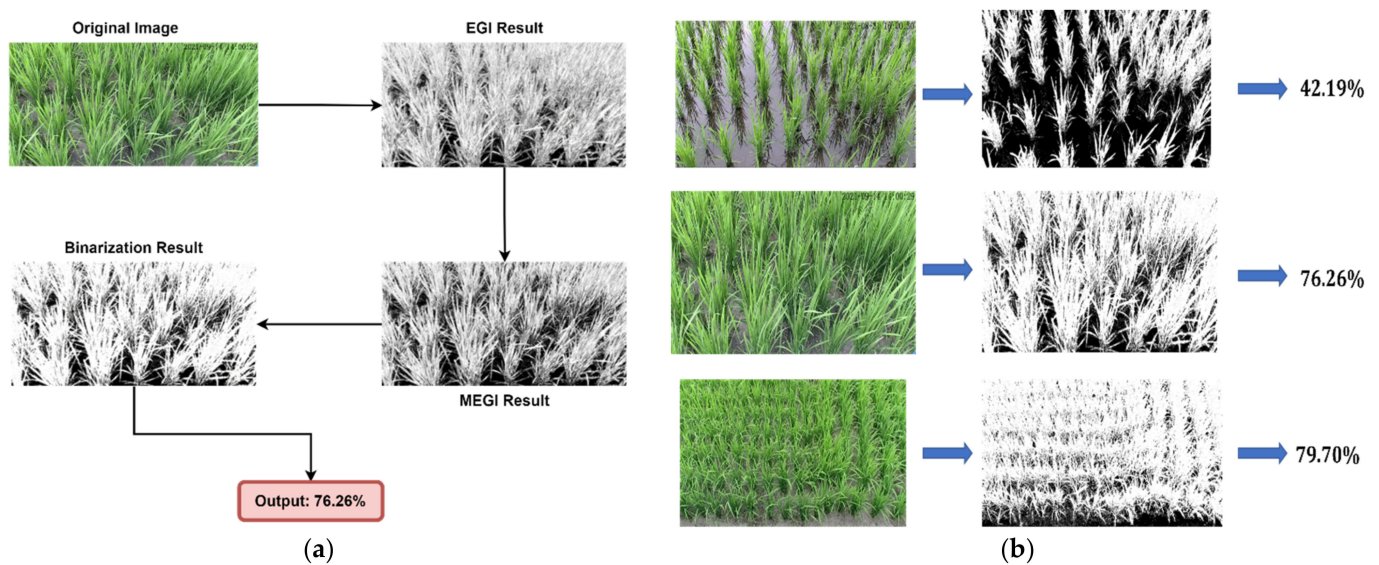
As images cannot depict straightened parts, error rates are still acceptable, and height is generally not a concern in later stages. Field 3 has a high error rate, so ground truth height is used in subsequent training instead of calibrated height.

#### 4.3. Green Coverage (GC) Rate

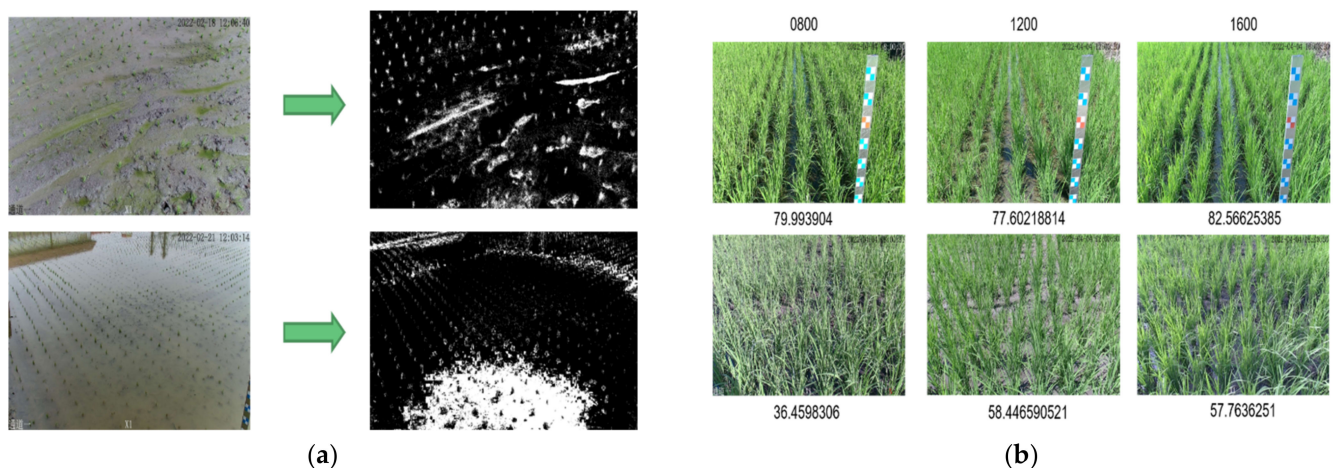
The workflow of GC calculation is shown in Figure 13a, as well as the results for each step. As we can see, EGI correctly extracts plant pixels from the background. MEGI modified the results to be more in line with the plants, with even more detail. Binarization has no difference with the MEGI results, but it is easier to calculate the ratio of the white pixel. Moreover, it was found that images with different capturing angles produce different results (cf. Figure 13b). In the future, it will be critical to develop and generalize the selection process of images to achieve better results.

Furthermore, it was found that algae-infested or flooded fields might not have accurate segmentation results, as shown in Figure 14a. The main reason is that algae or flooded fields have more G pixels than background pixels, causing the segmentation results to be less than ideal. It often happens in the early stages of the process, especially when transplantation is taking place. Since paddy seedlings cannot absorb fertilizer nutrients effectively, algae breed on them. Once those seedlings reach the tillering stage, these two problems are likely to reduce. In addition, the intensity of the lighting or sunlight affects the images as well. Due to the varying intensity of sunlight throughout the day (cf. Figure 14b), varying results are seen throughout the experiment time. Light reflection reduces the greenness of the paddy, leading to some extreme changes in the results.





**Figure 13.** GC Workflow and GC at Different Capturing Angles. (a) The workflow of the GC Procedure and the Output Results. (b) Different Results from Different Capturing Angles.

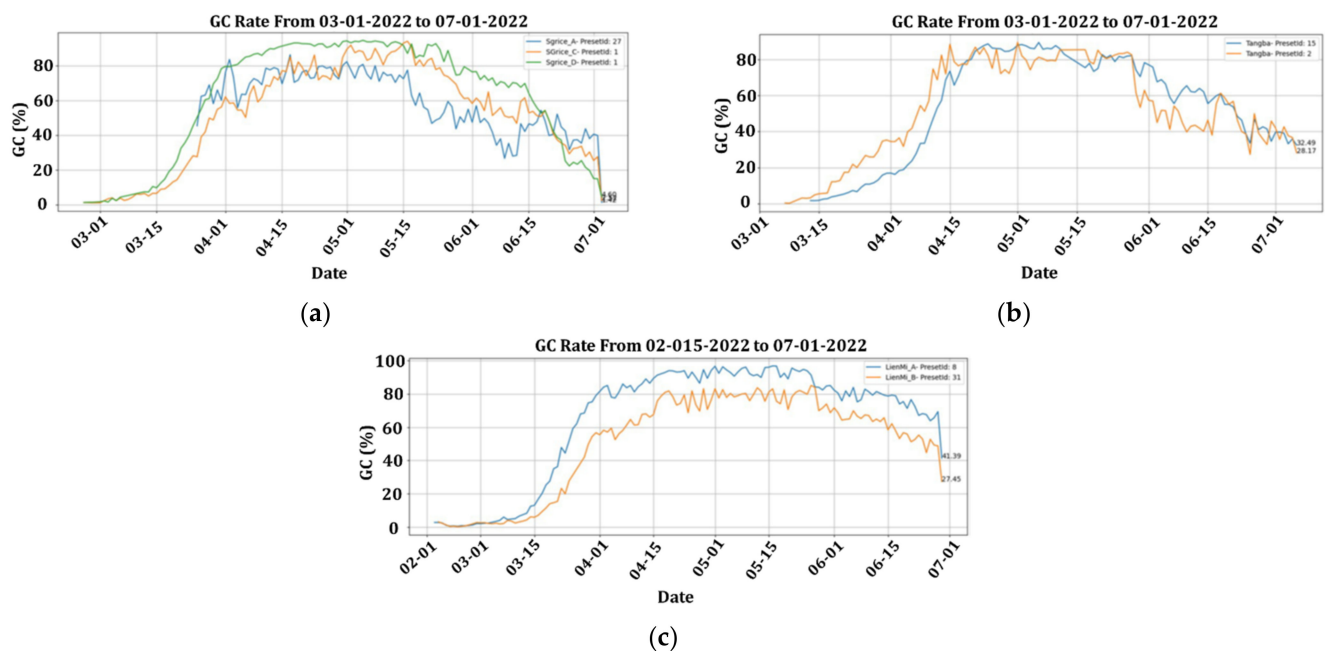


**Figure 14.** Effect of Algae and Light Intensity Throughout Day on the GC Rate. (a) Algae-infested and Flooded Fields. (b) Different Green Coverage Results Throughout the Day.

To reduce the deviation in the growth (tillering status), the results of each day must be averaged as shown in Figure 15. The GC results from three different fields were averaged and plotted in the graph below. In both cases, it was the first farming period in 2022. GC calculations are performed using at least two preset cameras per field.

By analyzing the graphs above, several phenomena can be observed. The GC rate tends to moderate after a rapid rise and gradually decreases over time. The trends were compared to the growth stages, and the summarized results are as follows: (a) GC rates tend to moderate after rapid increases, and then begin to decline after some time; (b) rapid growth could be considered the early (vegetative) stage, moderate growth (peak) would be the reproductive stage, and dropping growth would be the ripening stage; (c) the GC rate is reduced as the paddy plants start to turn yellow and panicles develop and mature. These two factors result in a decrease in green pixels; and (d) in each field, camera presets have a different GC rate due to the camera's height, angle, and captured images, but they tend to follow a similar trend.





**Figure 15.** GC Rate for the Three Experimental Fields. (a) GC in Field 1 having sgrice Rice. (b) GC in Field 2 having Tangba Rice. (c) GC in Field 2 having Tangba Rice.

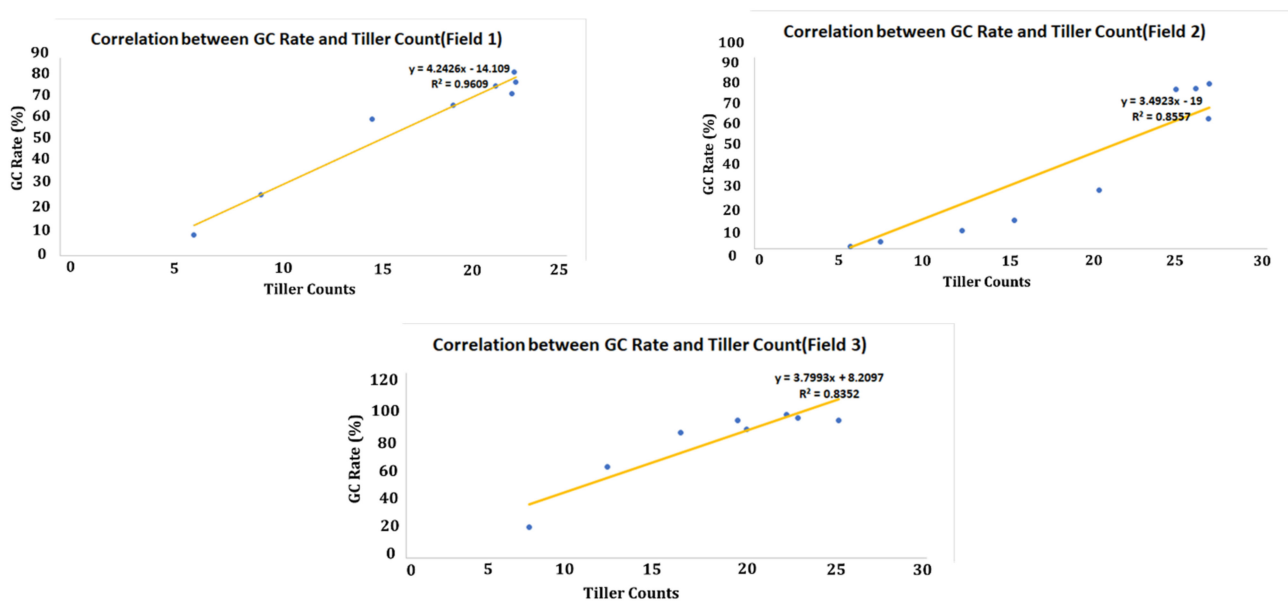
R-squared analysis was performed between GC rate and paddy tiller counts to determine the correlation between the two. For the calculation of the GC rate, we gathered the number of tillers on the paddy plants that were captured by the camera presets. Based on the range captured by the cameras, approximately eight to ten paddy instances were collected. In the following step, we averaged the numbers for each day. The tillering stage is the most representative stage of growth. Therefore, the stages before panicle initiation were focused on.

All of the R-squared results shown in Figure 16 are greater than 0.85, which is considered significant in statistical terms. This proves that the proposed method accurately measures the growth of paddy. As panicles mature, GC can still be used to determine maturity, even though growth is not the main factor. The green pixels and GC rate will decrease as the panicles and leaves gradually turn yellow later in the growth cycle.

#### 4.4. Performance Analysis of Classification Model

The dataset is split into a training dataset and a testing dataset in a ratio of 3:1, the number of tree estimators is set at 100, and the depth is freely expandable. For model performance evaluation, SciKit-Learn Library-based evaluation metrics are used. To evaluate the performance of the developed DL-based multi-class (different growth stages) classification model, macro metrics are employed. According to SciKit-Learn, macro metrics calculate metrics for each class and find their unweighted mean. This means that the imbalance problem will not be encountered, as each class will be treated equally. The SciKit-Learn Library treats accuracy as micro-averaging, which essentially computes the proportion of correctly classified data. Thus, to evaluate the model's performance, we mainly focus on the F1 score instead of accuracy. In addition, the F1-score is a better metric to evaluate a model in an imbalanced class distribution scenario, where accuracy can be misleading.

The baseline model produced an accuracy of 0.99454 and a macro F1-score of 0.97337, as shown in Table 11. The full metric results for each stage are presented in Table 12. From the results, it can be observed that the proposed baseline model performed well for all stages, except for Stage 1, which has a lower F1-score of 0.87500. A confusion matrix is used to describe the performance of a classification model on the test set and gives us more insight into the model's classification performance. As illustrated in Figure 17, the proposed baseline model has shown a high performance, resulting in very few negative results.



**Figure 16.** The R-Squared Analysis of Three Experimental Fields.

**Table 11.** The Performance of the RF Baseline Model.

	Accuracy	Macro F1-Score	Macro Precision	Macro Recall
Random Forest (baseline)	0.99454	0.97337	0.98883	0.96138

**Table 12.** The Performance Metrics of the RF Baseline Model for Each Stage.

Stage	Precision	Recall	F1-Score	Total
1	1.00000	0.77778	0.87500	9
2	0.94444	0.94444	0.94444	54
3	0.99563	0.99346	0.99454	917
4	0.99465	0.99785	0.99625	931
5	0.99666	0.99832	0.99749	597
6	0.99112	0.98674	0.98892	1131
7	0.98817	0.99331	0.99073	1345
8	1.00000	0.99915	0.99957	2347

The proposed machine-learning model's performance was further tested by utilizing the SHAP python-based library. As illustrated in Figure 18, the RF technique was used to retrieve the distribution of feature significance. Figure 18 indicates that accumulated temperature has the greatest value for feature importance followed by DAT and Height. Thus, these features are responsible for the better performance of the model.

To solve the imbalanced dataset problem, an up-sampling method was implemented. The RandomOverSampler (ROS) and SMOTE-ENN methods from the “imbalanced-learn” library 39 were employed to balance the dataset. RandomOverSampler (ROS) is an up-sampling method that over-samples the minority class(es) by picking samples at random with replacement. SMOTE-ENN is a combination of SMOTE 40 and ENN 41, in which SMOTE is used for up-sampling and ENN is used to clean the data to remove noise, overlapping, etc. Table 13 shows the data distribution after ROS and SMOTE-ENN were applied.

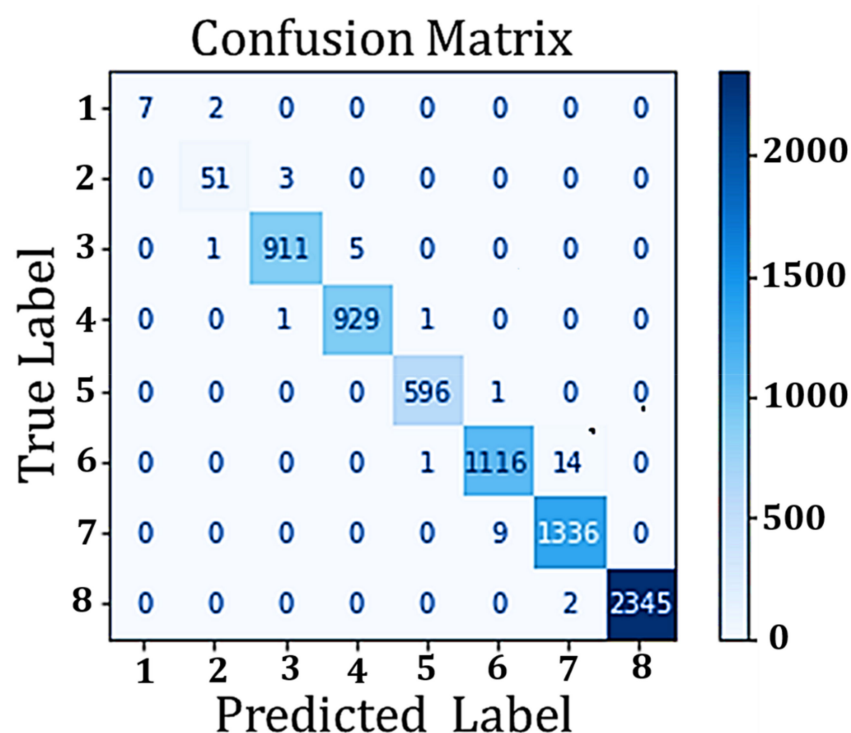


Figure 17. Confusion Matrix for Random Forest.

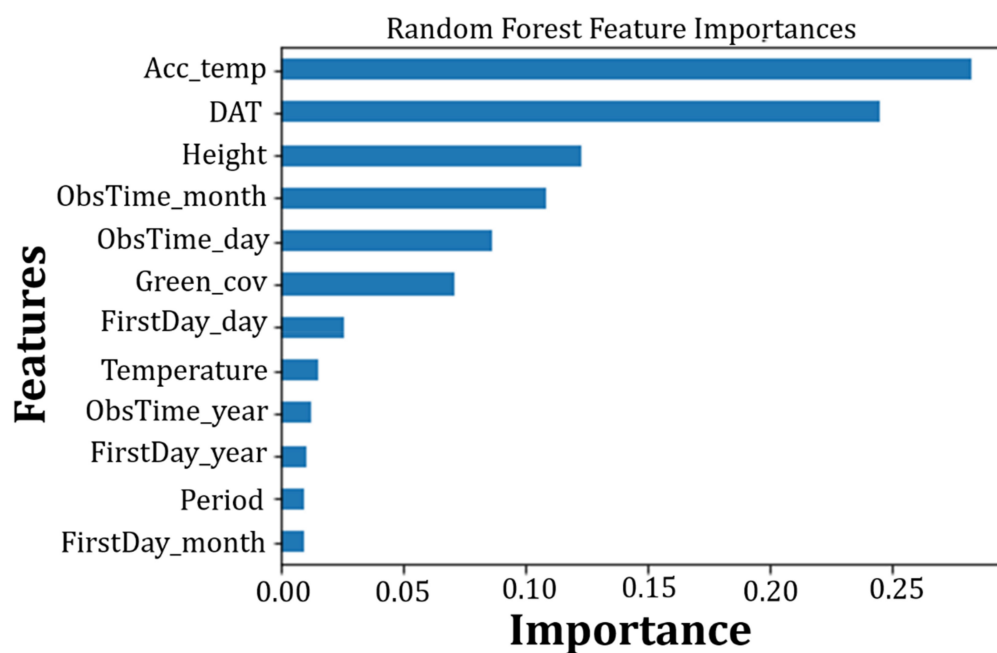


Figure 18. SHAP Library-The Feature Significance of the Proposed Model, Employing the RF.

Table 13. Training Dataset Before and After Up-sampling Methods.

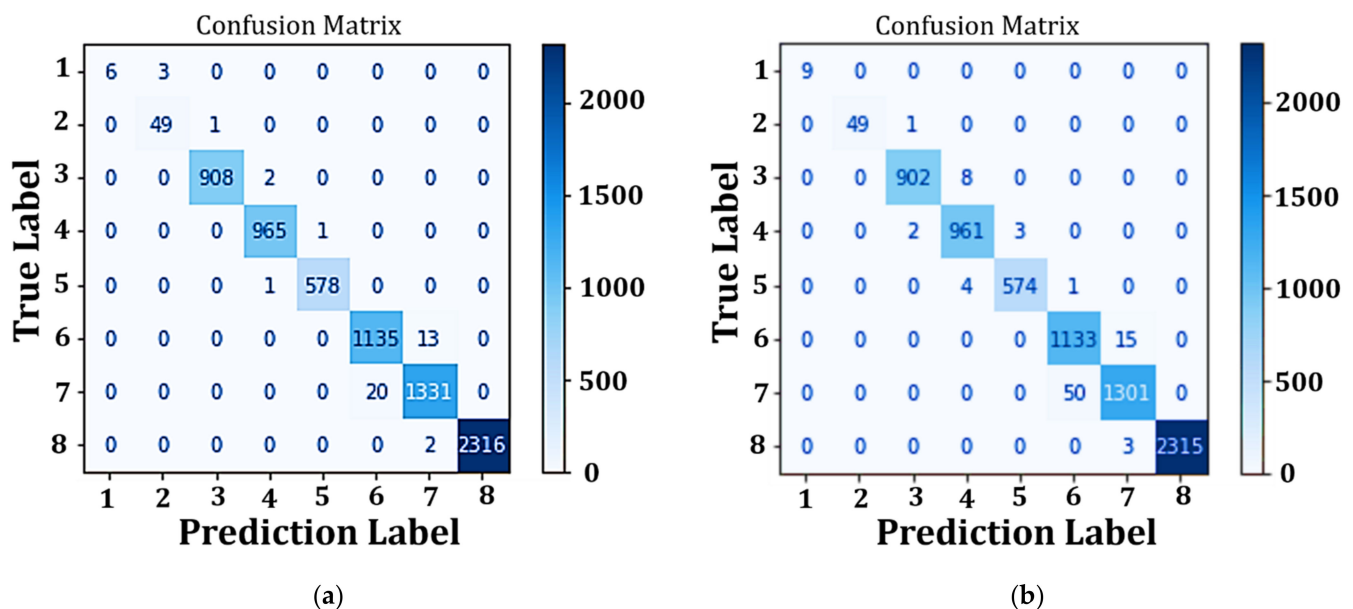
Stage	Before	RandomOverSampler	SMOTE-ENN	Stage	Before	RandomOverSampler	SMOTE-ENN
1	27	6953	6951	5	1737	6953	6932
2	150	6953	6923	6	3444	6953	6699
3	2728	6953	6922	7	4055	6953	6710
4	2899	6953	6884	8	6953	6953	6931

To build a generalized model with time-independent features, time-related features were dropped and the model was trained using leaf height, GC, DAT, and accumulated temperature, which is more growth-related. The results of the model trained after dropping certain features are shown in Table 14. The accuracy and the macro F1-score decreased slightly to 0.99018 and 0.95888, respectively. As a result, the proposed model does not rely on time-dependent features.

**Table 14.** Comparison After Up-sampling Methods Applied.

Model	Accuracy	Macro F1-Score	Macro Precision	Macro Recall
RF (Less Features)	0.99018	0.95888	0.96949	0.94932
RF (RandomOverSampler)	0.99373	0.96551	0.98805	0.95141
RF (SMOTE-ENN)	0.98772	0.98653	0.98518	0.98802

A comparison of the results of these two up-sampling methods was conducted to determine which performed better. A comparison of each method is shown in Table 14. ROS achieved the highest accuracy, but SMOTE-ENN had the highest F1-score. This is because ROS makes all other classes match the majority class, resulting in highly balanced data, which might have samples overlapping. The confusion matrix graphs for both methods are presented in Figure 19a,b. As can be seen from the confusion matrix graphs, both methods have a similar performance, but stage 1 is slightly different. ROS has a lower performance than SMOTE-ENN. SMOTE-ENN outperforms ROS in stages 1 and 2 (higher F1-Score of 1.00000 and 0.97030). Due to its highest F1-Score, SMOTE-ENN generally performs better, so for the remaining experiments, we use RF with SMOTE-ENN up-sampling.



**Figure 19.** Confusion Metrics for: (a) RF-(RandomOverSampler), and (b) RF-(SMOTE-ENN).

Additionally, we compared our model with existing machine learning models such as K-Nearest Neighbor (KNN), Support Vector Classifier (SVC), Adaptive Boost (AdaBoost), Gaussian Naive Bayes (GaussianNB), and Logistic Regression. These models are commonly employed in classification tasks and have shown good results in many instances. StandardScaler and SMOTE-EEN was used to compare performance, except for RF (only SMOTE-EEN is used). Scikit-Learn Library was used to implement all models, and their default parameters were used.



As can be seen from Table 15 shown above, the proposed model (RF with SMOTE-ENN) outperformed other models. SVC has the most promising results among the other five models, with an accuracy of 0.97313 and an F1-score of 0.96515. Followed by KNN, it has an accuracy of 0.96576 and a macro F1-score of 0.93469. Only SVC and KNN have their macro F1-score greater than 0.9, while others lie below 0.8 (inclusive).

**Table 15.** Classification Performance Comparison of Model with Existing ML Models.

Model	Accuracy	Macro F1-Score	Macro Precision	Macro Recall
RF (SMOTE-ENN)	0.98772	0.98653	0.98518	0.98802
KNN	0.96576	0.93469	0.95007	0.92945
SVC	0.97313	0.96515	0.95547	0.97744
AdaBoost	0.43609	0.32895	0.25549	0.47271
GaussianNB	0.83317	0.74009	0.71020	0.83983
Logistic Regression	0.86782	0.80107	0.76618	0.87496

The performance of GaussianNB was lower than expected. There was an assumption that there might be a zero-probability problem, which occurs when test data for a class are not included in training samples. Furthermore GaussianNB assumes that features are independent, but in real-world situations, features can be interdependent, for example, the height and GC rate change during DAT. As a result, Logistic Regression was also affected by this situation since it also requires that features be independent of one another. If one is related to the other, then the model will tend to emphasize the significance of those points.

AdaBoost had the worst results, with a macro F1-score of 0.32895. Due to AdaBoost's boosting technique (learns progressively), it is extremely crucial to ensure the quality of the data. Therefore, AdaBoost is extremely sensitive to noisy data and outliers. Almost all stages were misclassified by AdaBoost: these stages share some similarities with the previous or next stage, so AdaBoost had difficulty classifying them. Due to the aggregation of multiple tree outputs, RF remained robust to outliers and was less affected by noise. As a result, the developed model can be considered the most optimal choice for the particular classification task.

## 5. Conclusions and Future Works

In this paper, a classification model for paddy rice growth stages based on machine learning and image processing was designed and implemented. The developed system focuses on image data captured by the Speed-dome camera and growth factors determined from the collected images through image processing techniques (e.g., height, green coverage, etc.) installed in three different fields in Taiwan. The proposed system contains three components such as an Image Recognition Block, an Image Processing Block, and a Classification Model Block. Using an instance segmentation model, the first block will segment every paddy instance. The trained instance segmentation model has achieved a bounding box AP of 0.559, and segmentation AP of 0.506, which are both better than the Detectron2 baseline model. In addition, since not all instances were labeled, the Average Recall (AR) was taken into account. Both bounding box AR and segmentation AR are better than the baseline model, at 0.61 and 0.539, respectively.

The image processing block performs height and GC rate calculations. Every paddy instance predicted has its corresponding 2D image coordinates (bounding box coordinates) in the input image. By converting the 3D real-world object points/coordinates to 2D image coordinates, these coordinates are used to calculate the height of the paddy. Throughout the growth stages, it had a minimum error rate during the tillering stages and started increasing after the panicles were formed. It had an average error rate below 5%. Green Coverage sub-blocks were used to indicate vegetative status. The EGI/MEGI method was adopted, which is widely used in determining vegetative status or crop canopy cover. In

EGI, the green pixels are extracted from images to segment the plants from the background. As the color changes through the stages, the results can be used to represent growth status as well as the different growth stages.

The outputs of the image processing block were used for training the developed RF-based classification model. A paddy instance is considered a record of data. The proposed model was able to classify the growth stages with an accuracy of 0.99018 and an F1-score of 0.9588. Up-sampling methods were used to resolve the imbalanced dataset problem such as RandomOverSampler (ROS) and SMOTE-ENN. The SMOTE-ENN method produced the most accurate results. RF with SMOTE-ENN achieved an accuracy of 0.98772, and an F1-Score of 0.98653. With SMOTE-ENN applied, the performance in the classification of early stages improved significantly. The developed model was compared with other commonly used ML models. The proposed model with SMOTE-ENN outperformed all these models.

**Author Contributions:** Conceptualization, R.T.-C.S., Y.-H.H., P.-C.C., S.A.B., Y.-C.W. and N.-F.H.; methodology, S.A.B., R.T.-C.S., P.-C.C. and Y.-H.H.; validation, S.A.B., R.T.-C.S. and Y.-H.H.; formal analysis, S.A.B., R.T.-C.S., P.-C.C. and Y.-C.W.; investigation, S.A.B., R.T.-C.S. and N.-F.H.; resources, S.A.B. and N.-F.H.; data curation, R.T.-C.S., P.-C.C. and Y.-H.H., Y.-C.W. and N.-F.H.; writing—original draft preparation, S.A.B. and R.T.-C.S.; writing—review and editing, S.A.B. and N.-F.H.; visualization, R.T.-C.S., S.A.B., P.-C.C., Y.-H.H. and N.-F.H.; supervision, N.-F.H.; project administration, Y.-H.H., S.A.B., N.-F.H.; funding acquisition, N.-F.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the Ministry of Science and Technology (MOST 110-2622-8-816 007-009-TE2 and MOST 110-2923-E-007-008-) and the Ministry of Education (subsidy 817 for talent cultivation in international competition—College of Electrical Engineering and 818 Computer Science—5G-AIoT Technology and Application Research Center) in Taiwan.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Bhat, S.A.; Huang, N.-F.; Sofi, I.B.; Sultan, M. Agriculture-Food Supply Chain Management Based on Blockchain and IoT: A Narrative on Enterprise Blockchain Interoperability. *Agriculture* **2021**, *12*, 40. [CrossRef]
- Hsing, Y. Rice in taiwan. In *Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 3769–3772.
- Rice Can be Harvested Several Times a Year. Available online: [https://kmweb.coa.gov.tw/knowledge\\_view.php?id=167](https://kmweb.coa.gov.tw/knowledge_view.php?id=167) (accessed on 15 August 2022).
- Ramadhani, F.; Pullanagari, R.; Kereszturi, G.; Procter, J. Mapping a cloud-free rice growth stages using the integration of proba-v and sentinel-1 and its temporal correlation with sub-district statistics. *Remote Sens.* **2021**, *13*, 1498. [CrossRef]
- Production and Sales History Agricultural Products Production Process Taiwan Good Agricultural Practice (TGAP)-Rice-Paddy. Available online: <https://www.afa.gov.tw/cht/index.php> (accessed on 15 September 2022).
- Onyeneke, R.U.; Amadi, M.U.; Njoku, C.L.; Osuji, E.E. Climate Change Perception and Uptake of Climate-Smart Agriculture in Rice Production in Ebonyi State, Nigeria. *Atmosphere* **2021**, *12*, 1503. [CrossRef]
- Alfred, R.; Obit, J.H.; Chin, C.P.-Y.; Haviluddin, H.; Lim, Y. Towards paddy rice smart farming: A review on big data, machine learning, and rice production tasks. *IEEE Access* **2021**, *9*, 50358–50380. [CrossRef]
- Thakur, A.K.; Uphoff, N.T. How the system of rice intensification can contribute to climate-smart agriculture. *Agron. J.* **2017**, *109*, 1163–1182. [CrossRef]
- Bhat, S.A.; Huang, N.-F. Big data and ai revolution in precision agriculture: Survey and challenges. *IEEE Access* **2021**, *9*, 110209–110222. [CrossRef]
- Bhat, S.A.; Huang, N.-F.; Hussain, I.; Bibi, F.; Sajjad, U.; Sultan, M.; Alsubaie, A.S.; Mahmoud, K.H. On the Classification of a Greenhouse Environment for a Rose Crop Based on AI-Based Surrogate Models. *Sustainability* **2021**, *13*, 12166. [CrossRef]
- Huang, T.-W.; Bhat, S.A.; Huang, N.-F.; Chang, C.-Y.; Chan, P.-C.; Elepano, A.R. Artificial intelligence-based real-time pineapple quality classification using acoustic spectroscopy. *Agriculture* **2022**, *12*, 129. [CrossRef]

12. Muthusinghe, M.; Palliyaguru, S.; Weerakkody, W.; Saranga, A.H.; Rankothge, W. Towards smart farming: Accurate prediction of paddy harvest and rice demand. In Proceedings of the 2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Colombo, Sri Lanka, 6–8 December 2018; pp. 1–6.
13. Nishantha, M.D.L.C.; Zhao, X.; Jeewani, D.C.; Bian, J.; Nie, X.; Weining, S. Direct comparison of  $\beta$ -glucan content in wild and cultivated barley. *Int. J. Food Prop.* **2018**, *21*, 2218–2228. [\[CrossRef\]](#)
14. Vesali, F.; Omid, M.; Kaleita, A.; Mobli, H. Development of an android app to estimate chlorophyll content of corn leaves based on contact imaging. *Comput. Electron. Agric.* **2015**, *116*, 211–220. [\[CrossRef\]](#)
15. Haw, C.L.; Ismail, W.I.W.; Kairunniza-Bejo, S.; Putih, A.; Shamshiri, R. Colour vision to determine paddy maturity. *Int. J. Agric. Biol. Eng.* **2014**, *7*, 55–63.
16. Lee, K.-J.; Lee, B.-W. Estimating canopy cover from color digital camera image of rice field. *J. Crop Sci. Biotechnol.* **2011**, *14*, 151–155. [\[CrossRef\]](#)
17. Lee, K.-J.; Lee, B.-W. Estimation of rice growth and nitrogen nutrition status using color digital camera image analysis. *Eur. J. Agron.* **2013**, *48*, 57–65. [\[CrossRef\]](#)
18. Zainuddin, Z.; Manjang, S.; Wijaya, A.S. Rice farming age detection use drone based on SVM histogram image classification. In *Journal of Physics: Conference Series*; IOP Publishing: Bristol, UK, 2019; p. 092001.
19. Zhang, Y.; Xiao, D.; Liu, Y. Automatic Identification Algorithm of the Rice Tiller Period Based on PCA and SVM. *IEEE Access* **2021**, *9*, 86843–86854. [\[CrossRef\]](#)
20. Ikasari, I.H.; Ayumi, V.; Fanany, M.I.; Mulyono, S. Multiple regularizations deep learning for paddy growth stages classification from LANDSAT-8. In Proceedings of the 2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Malang, Indonesia, 15–16 October 2016; pp. 512–517.
21. Murata, K.; Ito, A.; Takahashi, Y.; Hatano, H. A study on growth stage classification of paddy rice by cnn using ndvi images. In Proceedings of the 2019 Cybersecurity and Cyberforensics Conference (CCC), Melbourne, Australia, 8–9 May 2019; pp. 85–90.
22. Wu, Y.; Kirillov, A.; Massa, F.; Lo, W.; Girshick, R. Detectron2 [WWW Document]. 2019. Available online: <https://github.com/facebookresearch/detectron2> (accessed on 3 March 2021).
23. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500.
24. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
25. Abdulla, W. Splash of color: Instance segmentation with mask r-cnn and tensorflow. *Matterport Eng. Techblog* **2018**. Available online: <https://engineering.matterport.com/splash-of-color-instance-segmentation-with-mask-r-cnn-and-tensorflow-7c761e238b46/> (accessed on 3 November 2022).
26. Nthu Smart Farming Platform. Available online: <https://nthu-smart-farming.kits.tw/> (accessed on 5 November 2022).
27. Paul, A.; Mukherjee, D.P.; Das, P.; Gangopadhyay, A.; Chintha, A.R.; Kundu, S. Improved random forest for classification. *IEEE Trans. Image Process.* **2018**, *27*, 4012–4024. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Chiu, M.-C.; Yan, W.-M.; Bhat, S.A.; Huang, N.-F. Development of smart aquaculture farm management system using IoT and AI-based surrogate models. *J. Agric. Food Res.* **2022**, *9*, 100357. [\[CrossRef\]](#)
29. Charbuty, B.; Abdulazeez, A. Classification based on decision tree algorithm for machine learning. *J. Appl. Sci. Technol. Trends* **2021**, *2*, 20–28. [\[CrossRef\]](#)
30. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2022.