

Low-cost and precise traditional Chinese medicinal tree pest and disease monitoring using UAV RGB image only

Haoran Guo

Hunan University of Chinese Medicine

Yuhua Cheng

Hunan University of Chinese Medicine

Jun Liu

`jun.liu@hnucm.edu.cn`

Hunan University of Chinese Medicine

Zhihu Wang

Hunan University of Chinese Medicine

Article

Keywords:

Posted Date: March 22nd, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4019957/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 26th, 2024.
See the published version at <https://doi.org/10.1038/s41598-024-76502-x>.

Low-cost and precise traditional Chinese medicinal tree pest and disease monitoring using UAV RGB image only

Haoran Guo¹, Yuhua Cheng¹, Jun Liu^{1,*} and Zhihu Wang¹

¹School of informatics, Hunan University of Chinese Medicine, Changsha, China;

*jun.liu@hnucm.edu.cn

ABSTRACT

Accurate and timely pest and disease monitoring during the cultivation process of traditional Chinese medicinal materials is crucial for ensuring optimal growth, increased yield, and enhanced content of effective components. This paper focuses on the essential requirements for pest and disease monitoring in a planting base of *Cinnamomum Camphora* var. *Borneol* (CCB) and presents a solution using unmanned aerial vehicle (UAV) images to address the limitations of real-time and on-site inspections. In contrast to existing solutions that rely on advanced sensors like multispectral or hyperspectral sensors mounted on UAVs, this paper utilizes visible light sensors directly. It introduces an ensemble learning approach for pest and disease monitoring of CCB trees based on RGB-derived vegetation indices and a combination of various machine learning algorithms. By leveraging the feature extraction capabilities of multiple algorithms such as RF, SVM, KNN, GBDT, XGBoost, GNB, and ELM, and incorporating morphological filtering post-processing and genetic algorithms to assign weights to each classifier for optimal weight combination, a novel ensemble learning strategy is proposed to significantly enhance the accuracy of pest and disease monitoring of CCB trees. Experimental results validate that the proposed method can achieve precise pest and disease monitoring with reduced training samples, exhibiting high generalization ability. It enables large-scale pest and disease monitoring at a low cost and high precision, thereby contributing to improved precision in the cultivation management of traditional Chinese medicinal materials.

Introduction

The *Cinnamomum Camphora* var. *Borneol* (CCB) belongs to the *Cinnamomum* genus in the Lauraceae family in plant taxonomy. It thrives in deep and loose soil, thrives in sunny and relatively warm and humid climates, but is sensitive to high temperatures. The CCB tree, along with its abundant natural dextro-rotatory camphor [1], also known as natural borneol, possess remarkable medicinal properties [2]. Not only is it a rare and precious medicinal herb, but it also serves as a premium spice and a vital chemical raw material. Due to its exceptionally limited natural distribution, estimated at only about one in ten thousand, the camphor tree is revered as "plant gold" for its scarcity and high value.

In 1988, the first wild CCB tree rich in borneol was discovered in the pristine forest of Butoujiang Miao Township, Xinhuan Dong Autonomous County, Hunan. Today, extensive and standardized CCB tree plantations have been established in Xinhuan, Hunan, and Ji'an, Jiangxi. Current research primarily focuses on the volatile oil extracted from CCB trees. Experimental studies have revealed that the highest concentration of dextrorotatory borneol (natural Borneo camphor) is present in the fresh leaves dried in the shade among various parts of the CCB tree (such as fresh leaves dried in shade, fresh leaves, fresh branches, fresh twigs, fallen leaves, etc.) [3]. The large-scale and intensive cultivation of CCB trees is particularly vulnerable to pest and disease outbreaks. Without prompt preventive measures, these issues could result in substantial losses for the CCB tree industry and associated medicinal practitioners. Therefore, timely and precise monitoring of pests and diseases affecting CCB trees is essential to safeguarding their economic value.

Remote sensing technology stands out as one of the most cutting-edge methods globally for monitoring pests and diseases. Back in the 1930s and 1940s, remote sensing was already being utilized to keep track of forest pests and diseases. Fast forward to the 21st century, with the rapid advancements in space technology, nations have poured substantial resources into the research and development of satellite remote sensing technology. This investment has led to the launch of various satellite platforms, resulting in significant enhancements in remote sensing image data quality in terms of spatial, temporal, spectral, and radiometric resolution. Simultaneously, remote sensing imaging techniques have diversified, with Light Detection and Ranging (LiDAR) and Search and Rescue (SAR) platforms making notable progress. Notably, both manned and unmanned aerial sensors have streamlined data collection, providing an unprecedented wealth of data for monitoring pests and diseases in forestry. The primary methods [4] for extracting pest and disease information

encompass: (1) image classification; (2) a range of vegetation indices and ratios; (3) differencing techniques; (4) image processing algorithms; (5) spectral analysis methodologies; and (6) mathematical statistics approaches coupled with GIS technology. Researchers have leveraged vegetation index time series data to investigate forest defoliation caused by pests and diseases like the gypsy moth and the pine processionary moth in countries such as the United States, Sweden, and China, achieving commendable monitoring precision [5].

Remote sensing satellites offer a wide coverage area, yet they come with drawbacks such as limited data sources, delayed data delivery, and lower resolution. The advancement of small civilian unmanned aerial vehicle (UAV) technology has addressed these limitations of satellite-based remote sensing. By leveraging UAVs as platforms for remote sensing sensors, new avenues have opened up for the exploration of forest pest and disease monitoring through remote sensing. Multi-spectral/hyperspectral remote sensing on UAVs presents a more dependable solution for numerous agricultural crops that require enhanced spectral resolution and richer vegetation details. Xu's research primarily delved into the impact of the major leaf-eating pest *Pantana phyllostachysae* Chao (PPC) on moso bamboo. Damage-sensitive information was extracted from multi-spectral images captured by UAVs, which were then used to train machine learning algorithms such as random forest (RF), extreme gradient boosting tree (XGBoost), and support vector machine (SVM). The findings indicated that XGBoost exhibited the most superior detection performance, followed by RF and SVM [6]. In addition to multi-spectral imaging, other researchers have utilized UAVs equipped with hyperspectral sensors to conduct remote sensing studies on pests and diseases, achieving remarkable accuracy. Jiang et al [7]. employed the first-order derivative maximum method to identify the spectral red edge and yellow edge positions for constructing a vegetation index, enabling the early detection of wheat diseases up to 12 days in advance with high precision. Nevertheless, the demands placed on multi-spectral/hyperspectral remote sensing sensors on UAVs are rigorous, and practical applications may encounter certain limitations.

Drawing upon the response mechanism of healthy green vegetation to the visible light spectrum, we understand that green plants absorb blue and red light while reflecting green light. UAVs equipped with RGB cameras can swiftly capture grayscale or color images within the visible light spectrum, along with vegetation height maps. This capability facilitates the generation of diverse visible light vegetation indices, enabling the estimation of crop height, leaf angle distribution, leaf area index, lodging area, leaf color, biomass, and other critical growth parameters [8]. Current research on UAV RGB imagery can be broadly classified into two primary approaches. The first approach involves computing pertinent vegetation indices for crops using UAV RGB images and analyzing physiological indicators and vegetation indices via conventional machine learning models. This method allows for the rapid assessment of crop growth patterns and disease-related information. On the other hand, the second approach entails leveraging local image processing techniques, such as deep learning and semantic segmentation, to analyze complete UAV RGB aerial images comprehensively.

In the realm of vegetation information extraction, Kazmi et al. [9] employed 14 visible light vegetation indices, including EXG, CIVE, and GB, to identify thistles in beetroot fields. For vegetation biomass estimation, a comprehensive method [10] utilizing UAV images was developed to assess the uniformity of wheat emergence rate. This method combined the EXG index and OTSU method to analyze wheat seedling coverage. Zhang et al. [11] conducted an analysis on the relationship between nine visible light vegetation indices and vegetation coverage across four crops (sugarcane, corn, cotton, and rice) using statistical methods. Their findings revealed that the accuracy of the vegetation coverage model and the optimal vegetation index varied among different crop types. Görlich et al. delved into the detection of white mold (*Cercospora betticola*) induced by beetroot leaf spot disease (CLS) using RGB images. They proposed a method employing a fully convolutional neural network to address the CLS detection challenge. This approach directly processes RGB images captured by UAVs, achieving an impressive F1 score of 88% for CLS recognition [12]. Furthermore, when crops exhibit deficiencies in essential elements leading to reduced chlorophyll levels and notable alterations in leaf color, UAV RGB technology proves effective in detecting such conditions on a large scale. For instance, Zermas et al. captured high-resolution UAV RGB images of corn fields and utilized support vector machines (SVM) in machine learning to identify nitrogen-deficient corn, achieving an average accuracy of 82.3% [13] in detecting nitrogen-deficient leaves. These studies underscore the efficacy of RGB-based vegetation indices as a valuable technical tool for vegetation monitoring, as well as for the surveillance of vegetable pests and diseases.

Machine learning, especially deep learning, is a research field dedicated to enhancing computer algorithm performance through experiential learning, by generating models from vast amounts of data [14]. Deep learning has seen widespread application in recent years, particularly in the realm of agricultural pest and disease detection and recognition. Yi Yue et al. [15] introduced a super-resolution model to tackle the resolution challenges of agricultural pest images captured by cameras. Pan Chunhua et al. [16] proposed a methodology leveraging region growth and SVM algorithms to classify and identify four major pests common in South China's vegetables (striped flea beetles, aphids, cabbage moths, and thrips) with an impressive identification accuracy of up to 93%. Xi Cheng et al. [17] utilized deep learning techniques to address the complexities associated with identifying agricultural pests in intricate environments. Sharada P. Mohanty et al. [18] trained a deep convolutional neural network to distinguish among 14 crops and 26 crop diseases, achieving an overall accuracy

ranging from 85.53% (using the AlexNet transfer model) to 99.34% (employing the GoogleNet transfer model) on the PlantVillage dataset.

In various classification scenarios and vegetation types, the generalization of a single machine learning algorithm may exhibit instability. Ensemble learning (EL) algorithms offer a solution by amalgamating multiple machine learning classifiers to create models with enhanced generalization capabilities and superior resistance to overfitting [19]. Scholars have developed a stacking ensemble model for inverting water quality parameters using a combination of ML, RF, BP, and XGBoost. This ensemble model outperforms individual models, with the top inversion model achieving an R^2 value of 0.839, signifying significant potential for water quality monitoring [20]. Furthermore, researchers have devised an ensemble learning model by stacking five fundamental models (AdaBoost, XGBoost, CatBoost, LightGBM, and Random Forest), showcasing that the ensemble model's generalization prowess surpasses that of a standard single model [21].

In summary, UAV remote sensing presents notable advantages for large-scale planting bases characterized by intricate terrains and dense forests where conducting on-site investigations proves challenging. While multispectral/hyperspectral images entail high acquisition costs and processing complexities, visible light images offer the benefits of cost-efficiency and operational ease. Multi-model ensembles frequently exhibit superior classification capabilities compared to individual models; however, the efficacy of ensemble models hinges on the classification proficiency of the underlying base models and the ensemble strategy employed. Therefore, this study introduces a refined ensemble strategy leveraging visible light images and genetic algorithms, which substantially enhances the accuracy of pest and disease monitoring in CCB trees.

The primary contributions encompass:

- . In contrast to the traditional approach of relying solely on RGB bands and a single vegetation index, this paper introduces a novel method that leverages multiple vegetation indices to extract diverse vegetation feature information across various dimensions. This innovative approach leads to enhanced pest and disease monitoring capabilities;

- . By employing an ensemble learning technique that incorporates a range of vegetation index feature information, this study integrates the information extraction capabilities of multiple machine learning algorithms, including RF, SVM, KNN, GBDT, XGBoost, GNB, and ELM. This comprehensive approach enhances the expression of pest and disease information features and ultimately results in improved monitoring accuracy.

- . This paper sets itself apart from conventional approaches by considering the varying information extraction capabilities and emphases of different machine learning algorithms. It introduces a genetic algorithm to assign optimal weights to each algorithm, creating a novel majority voting strategy. Empirical findings confirm that this innovative strategy substantially enhances the accuracy of pest and disease monitoring;

- . To tackle the challenge of reduced monitoring accuracy due to noise in small-sample conditions, this paper proposes a morphological post-processing method. When integrated with an ensemble learning voting strategy, this approach can notably enhance accuracy in scenarios with limited data samples;

- . With the proposed method, cost-effective, efficient, and precise pest and disease monitoring for CCB tree planting bases can be realized. Across diverse sample conditions, the monitoring accuracy consistently outperforms that of individual machine learning algorithms, effectively mitigating noise and showcasing strong generalization capabilities.

Results

By leveraging the genetic algorithm-based ensemble strategy introduced in this paper, pests and diseases within the CCB planting base are effectively monitored and identified through the utilization of data gathered from UAVs. The algorithm's efficacy is assessed through the application of various accuracy metrics to gauge its performance.

Parameter settings

Utilizing RGB-derived vegetation indices, a training dataset comprising 50 randomly selected samples from both healthy and diseased plants was created. Detection result images were generated by employing seven machine learning classifiers, along with direct majority voting ensemble strategy and genetic algorithm ensemble strategy. Various accuracy metrics were computed based on a confusion matrix. Furthermore, to evaluate the effectiveness of utilizing vegetation indices versus original RGB data, the same samples, methods, and ensemble strategies were applied to RGB images, and the accuracy of the detection results was assessed for comparison.

All experiments were carried out in Python 3.6, with the machine learning algorithms implemented using scikit-learn version 1.3.1.

The parameter settings for the seven machine learning classifiers are listed below:

- . RF: The n_estimators is set to 200, and all other parameters are default.
- . SVM: The classifier uses all default parameters.
- . KNN: The n_neighbors is set to 3, and all other parameters are default.
- . GBDT: The random_state is set to 37, and all other parameters are default.
- . XGBoost: The n_estimators is set to 100, and all other parameters are default.
- . GNB: The classifier uses all default parameters.
- . ELM: The classifier uses all default parameters.

Accuracy Assessment

The confusion matrix is a widely used representation for evaluating the accuracy of machine learning classification tasks. It provides a visual and detailed overview of an algorithm's performance in classification tasks. In binary classification scenarios analyzed using machine learning models, a 2x2 confusion matrix is generated, as illustrated in Table 1. This matrix allows for the derivation of various accuracy metrics to evaluate the effectiveness of different classification approaches and to assess the quality of the classification model and its outcomes.

By comparing the actual ground truth with the predicted labels, where 1 denotes trees affected by pest and disease damage, and 0 denotes healthy trees, the test samples can be categorized into four groups based on the accuracy of the predictions: A represents the pixels correctly identified as pest-damaged trees, B represents the pixels incorrectly classified as pest-damaged when they are actually healthy trees, C represents the pixels incorrectly classified as healthy when they are actually pest-damaged trees, and D represents the pixels correctly identified as healthy trees. These classification scenarios allow for the calculation of various metrics to assess the effectiveness of the machine learning model.

	Prediction value is 1	Prediction value is 0
The true value is 1	A	B
The true value is 0	C	D

Table 1. Confusion Matrix

The Critical Success Index is defined as the ratio of samples in which the model predicts a positive outcome and the actual value is also positive, divided by the total number of samples, expressed by the following formula:

$$CSI = \frac{A}{A+B+C}, \quad (1)$$

The Probability of Detection refers to the ratio of samples in which the model predicts a positive outcome and the actual value is also positive, among all positive samples, calculated using the following formula:

$$POD = \frac{A}{A+B}, \quad (2)$$

The False Alarm Rate represents the percentage of samples predicted as positive by the model but are actually negative, out of all the samples predicted as positive. It can be calculated using the following formula:

$$FAR = \frac{C}{A+C}, \quad (3)$$

Overall Accuracy (OA) is the likelihood that the classification result of each random sample aligns with the test data type. It is commonly used alongside the Kappa coefficient. The calculation for Overall Accuracy is as follows:

$$OA = \frac{A+D}{A+B+C+D}, \quad (4)$$

The Kappa coefficient is a metric used to gauge the consistency and efficacy of classification testing. It evaluates the agreement between predicted outcomes and actual classifications in classification scenarios. Derived from the confusion matrix, the Kappa coefficient usually falls within the range of -1 to 1, with values above 0 indicating agreement. The calculation formula is provided below:

$$\text{kappa} = \frac{P_o - P_e}{1 - P_e} \quad (5)$$

where, $P_o = \frac{\text{The sum of the diagonal elements}}{\text{The sum of all matrix elements}}$;

$$P_e = \frac{\sum_i \text{The sum of elements in the } i\text{-th row} * \text{The sum of elements in the } i\text{-th column}}{\left(\sum \text{All elements of the matrix} \right)^2}$$

, that is, the sum of the products of the actual and predicted numbers for all categories, divided by the square of the total number of samples.

Results without morphological post-processing

The pest and disease monitoring results, based on the original RGB image and RGB-derived vegetation indices, are presented in Figures 1 and 2, respectively. The white areas in the figures indicate the detected pest and disease plants without morphological post-processing. Overall, all methods demonstrate effective detection of most pest and disease trees when compared to the ground truth, utilizing both data sources. However, the classification outcomes derived from the original RGB image display varying levels of noise, whereas those from the RGB-derived vegetation indices exhibit reduced noise and clearer image boundaries. Employing the direct majority voting ensemble strategy and the ensemble strategy based on genetic algorithms helps mitigate noise to some extent. The classification results using vegetation indices more accurately depict the actual positions and distribution of pest and disease trees in the UAV images.

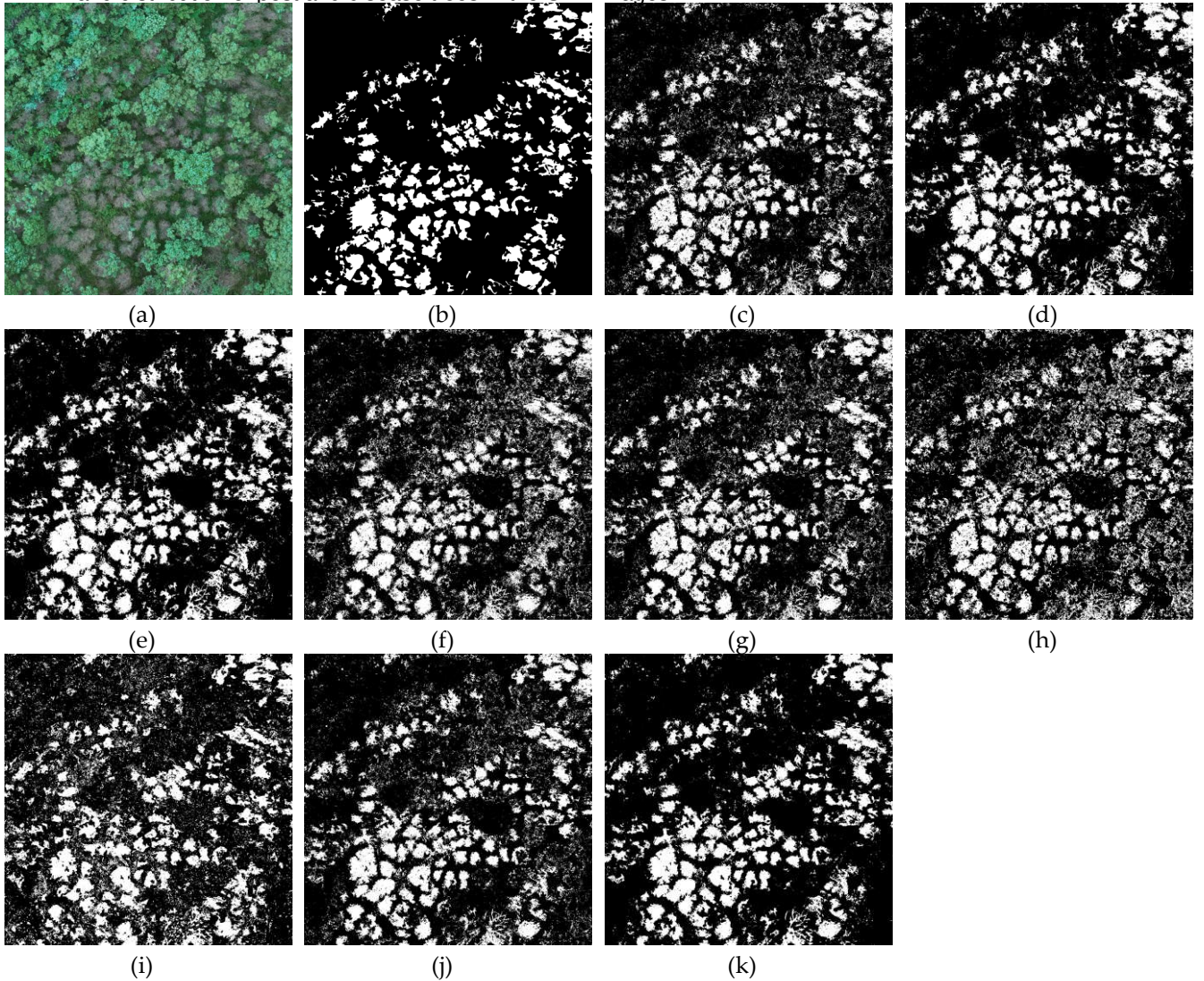


Figure 1 . The classification results without morphological post-processing based on original RGB images. (a) Original RGB image; (b) Ground truth; (c) RF; (d) SVM; (e) KNN; (f) GBDT; (g) XGBoost; (h) GNB; (i) ELM; (j) MVEs; (k) GAEns.

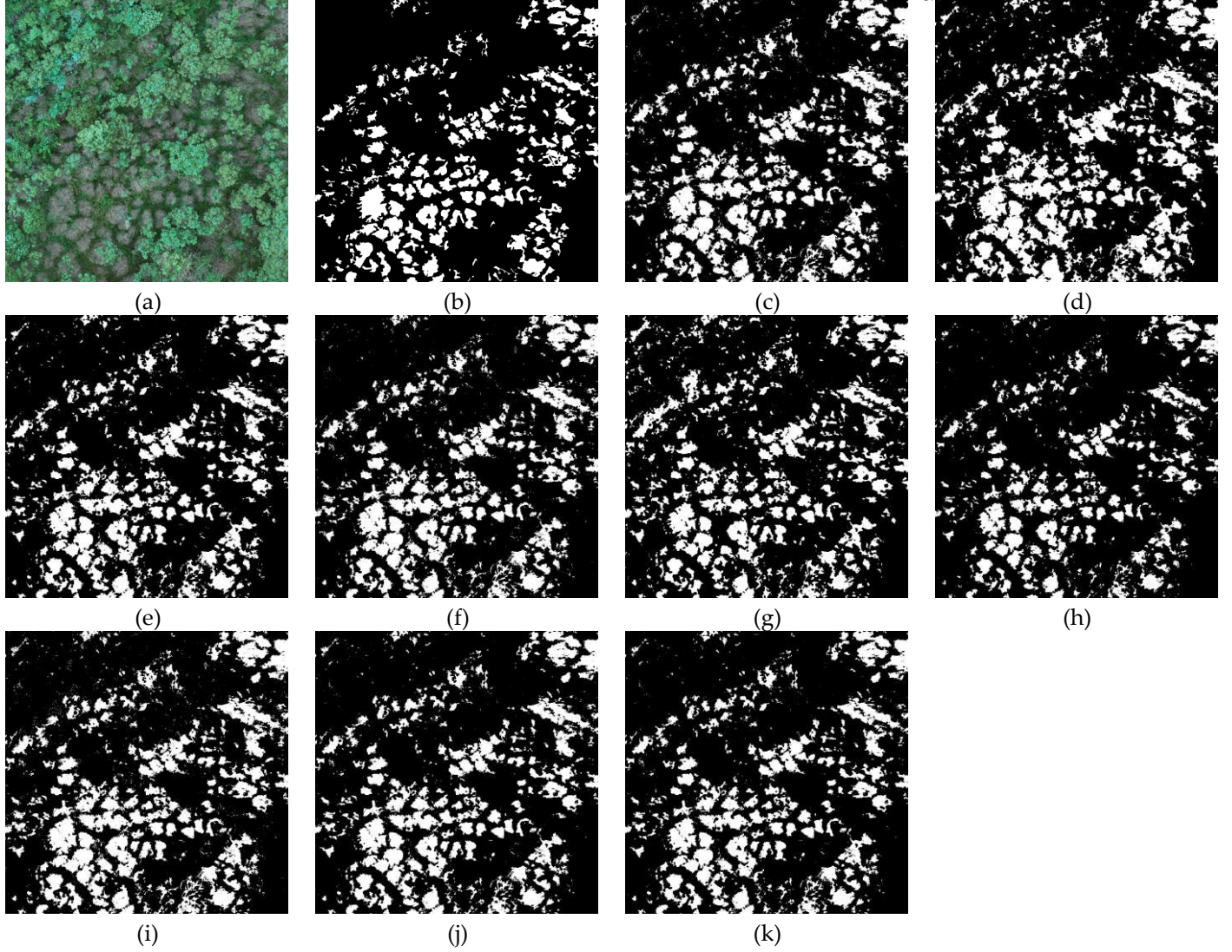


Figure 2 . The classification results without morphological post-processing based on the RGB-derived vegetation indices. (a) Original RGB image; (b) Ground truth; (c) RF; (d) SVM; (e) KNN; (f) GBDT; (g) XGBoost; (h) GNB; (i) ELM; (j) MVEs; (k) GAEns.

Various accuracy evaluation metrics have been computed for the aforementioned monitoring results, as outlined in Table 2. Across most machine learning methods, it is evident that the outcomes derived from the RGB-derived vegetation indices outperform those from the original RGB images across all metrics. The presence of noise emerges as a crucial factor contributing to this observation. While the accuracy of the direct majority voting ensemble strategy may not surpass that of individual methods like SVM and KNN, the accuracy of the genetic algorithm ensemble strategy exhibits enhancements over all individual machine learning techniques. For instance, when utilizing the original RGB image as the data source, the CSI metric of GNB has seen an improvement from 0.3596 to 0.6135, with OA also advancing from 0.7877 to 0.9047. Furthermore, the OA achieved when utilizing vegetation indices as the data source reached 0.9233, highlighting the benefits of ensemble learning and genetic algorithm optimization.

	CSI		POD		FAR		OA		Kappa	
	RGB	VI	RGB	VI	RGB	VI	RGB	VI	RGB	VI
RF	0.4882	0.6590	0.7417	0.8789	0.4119	0.2752	0.8606	0.9185	0.5701	0.7442
SVM	0.5853	0.6040	0.8289	0.9177	0.3342	0.3615	0.8947	0.8921	0.6735	0.6869
KNN	0.5763	0.6508	0.8727	0.8567	0.3708	0.2697	0.8850	0.9176	0.6604	0.7377
GBDT	0.4548	0.6597	0.7619	0.8700	0.4698	0.2682	0.8363	0.9195	0.5248	0.7453
XGBoost	0.5038	0.5845	0.7772	0.8197	0.4111	0.3293	0.8628	0.8955	0.5855	0.6733

GNB	0.3596	0.6411	0.6651	0.7359	0.5608	0.1673	0.7877	0.9261	0.3993	0.7371
ELM	0.4285	0.6369	0.7542	0.8993	0.5020	0.3142	0.8196	0.9081	0.4897	0.7215
MVEns	0.5116	0.6652	0.7741	0.8732	0.3986	0.2637	0.8675	0.9212	0.5952	0.7504
GAEns	0.6135	0.6674	0.8440	0.8583	0.3080	0.2500	0.9047	0.9233	0.7017	0.7533

Table 2. Classification accuracy without morphological post-processing

Results with morphological post-processing

Following the implementation of morphological post-processing, the detection outcomes from the two data sources are visually presented in Figures 3 and 4. Upon inspecting the detection results obtained from the original RGB images, it is evident that, in comparison to Figure 1, the noise associated with each machine learning method undergoes a significant reduction post morphological post-processing. Additionally, the noise level in the ensemble result is also notably diminished. Similarly, the detection results derived from the vegetation indices exhibit a lower noise level than Figure 2. However, the extent of noise reduction is comparatively less pronounced than that observed in the original RGB images, suggesting that the vegetation indices possess the capability to mitigate noise generation to a certain degree.

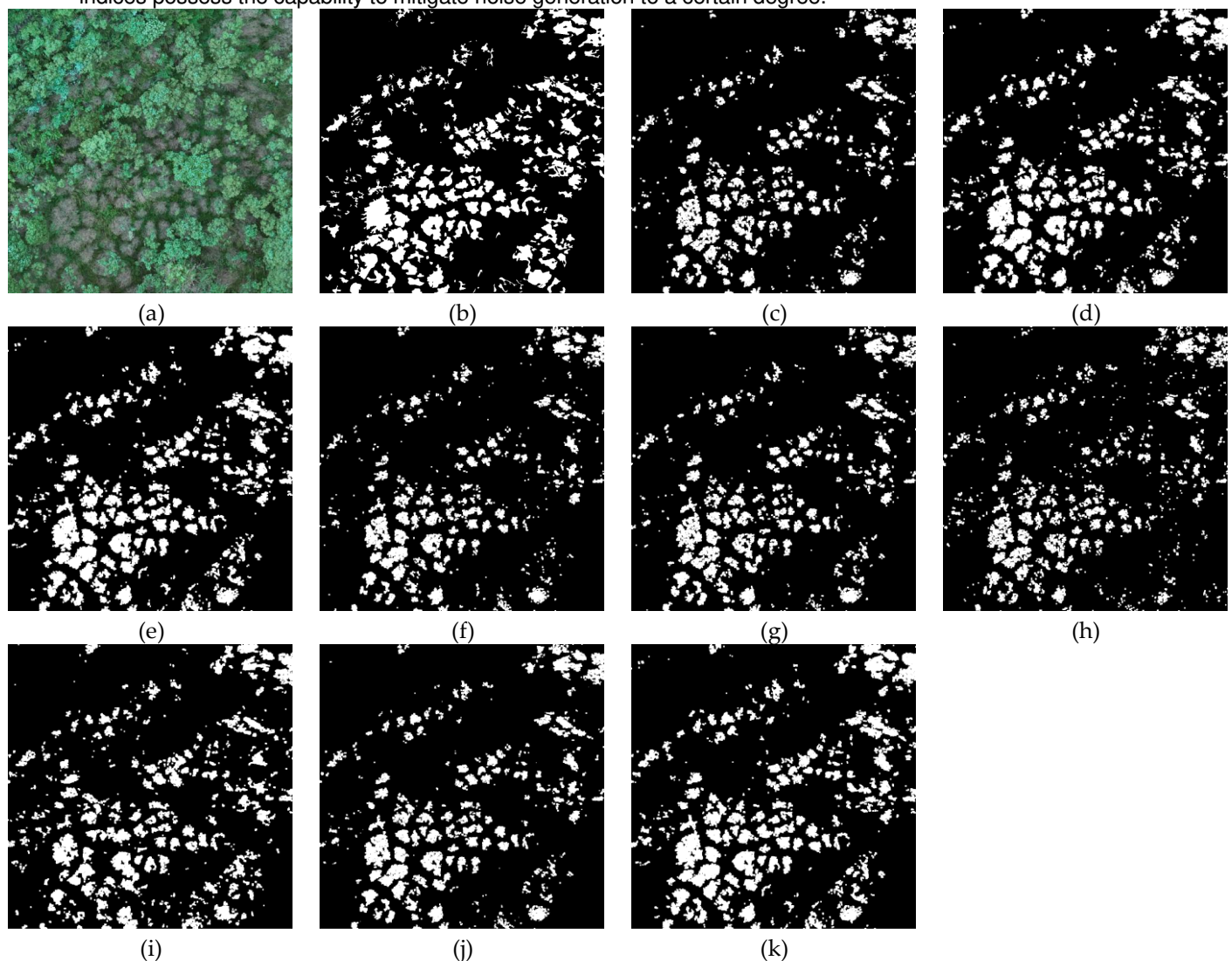


Figure 3. Classification results based on original RGB image with morphological post-processing. (a) Original image; (b) Ground truth; (c) RF; (d) SVM; (e) KNN; (f) GBDT; (g) XGBoost; (h) GNB; (i) ELM; (j) MVEns; (k) GAEns.

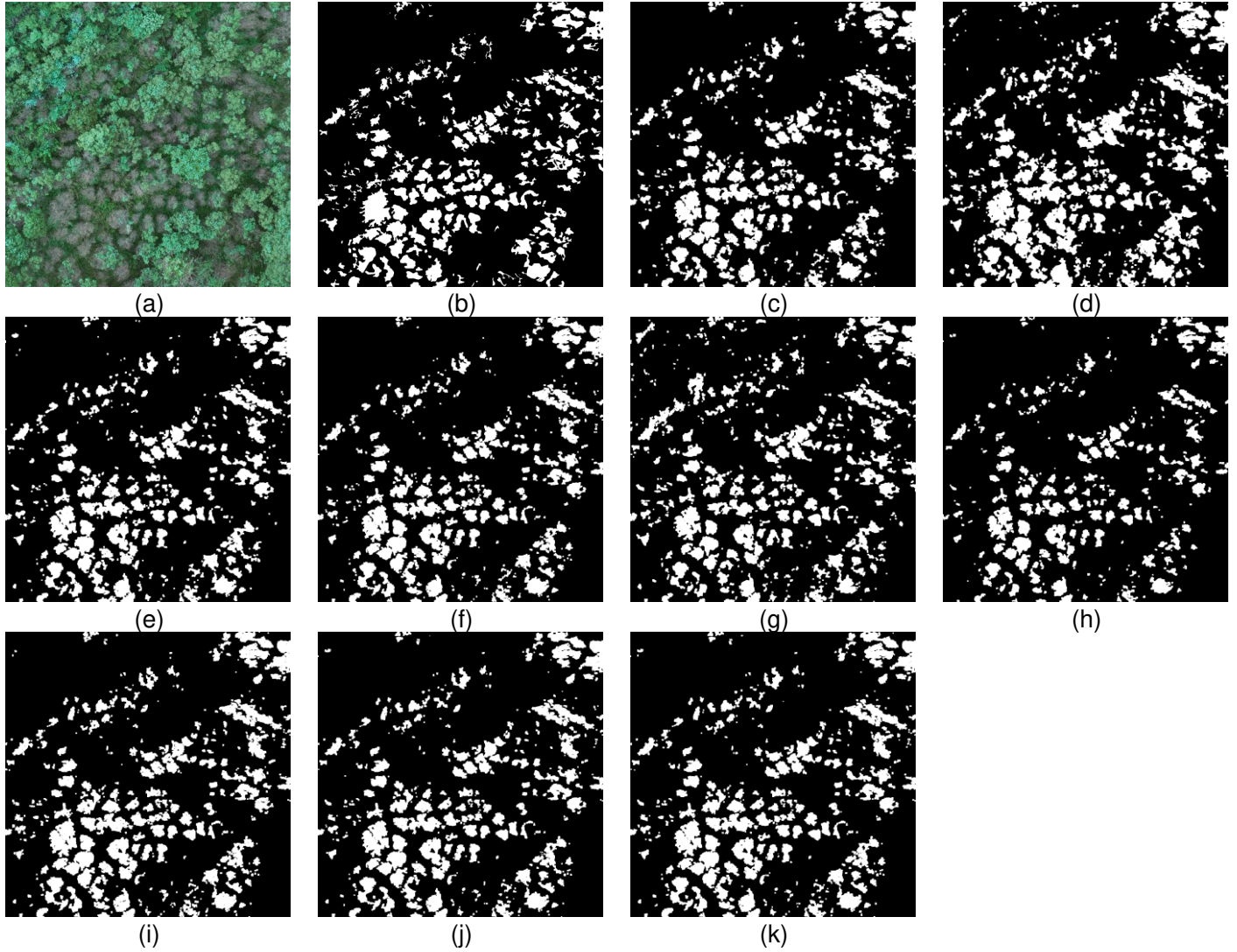


Figure 4. Classification results based on the RGB-derived vegetation indices with morphological post-processing. (a) Original image; (b) Ground truth; (c) RF; (d) SVM; (e) KNN; (f) GBDT; (g) XGBoost; (h) GNB; (i) ELM; (j) MVEEns; (k) GAEns.

The quantitative accuracy indicators' computation results are detailed in Table 3. A comparison with Table 2 reveals a notable enhancement in the accuracy of each machine learning method when utilizing the original RGB image as the data source. For instance, the CSI of KNN has surged from 0.5763 to 0.6306, with the OA also increasing from 0.8850 to 0.9211. Conversely, the accuracy of the outcomes derived from the vegetation indices displays only marginal improvements. Furthermore, the discrepancy between Figure 4 and Figure 2 suggests that the influence of noise is less pronounced compared to the variance between Figure 3 and Figure 1. Analogous to the findings in Table 2, the accuracy of direct majority voting ensemble does not consistently surpass that of individual machine learning algorithms, whereas the accuracy of the ensemble strategy predicated on genetic algorithm optimization outperforms that of individual machine learning algorithms.

	CSI		POD		FAR		OA		Kappa	
	RGB	VI	RGB	VI	RGB	VI	RGB	VI	RGB	VI
RF	0.4972	0.6819	0.5304	0.8137	0.1119	0.1920	0.9038	0.9319	0.6122	0.7693
SVM	0.6034	0.6387	0.6943	0.8864	0.1784	0.3043	0.9182	0.9101	0.7040	0.7241
KNN	0.6306	0.6606	0.7510	0.7955	0.2026	0.2043	0.9211	0.9267	0.7258	0.7510
GBDT	0.4657	0.6731	0.5009	0.7912	0.1309	0.1815	0.8970	0.9311	0.5805	0.7628
XGBoost	0.5304	0.6026	0.5704	0.7769	0.1168	0.2712	0.9095	0.9082	0.6429	0.6958
GNB	0.3997	0.5967	0.4509	0.6438	0.2212	0.1094	0.8786	0.9220	0.5062	0.7027
ELM	0.5367	0.6814	0.6401	0.8351	0.2314	0.2126	0.9009	0.9300	0.6398	0.7676
MVEEns	0.5423	0.6825	0.5886	0.8097	0.1268	0.1871	0.9109	0.9325	0.6531	0.7702
GAEns	0.6308	0.6861	0.7514	0.8264	0.2029	0.1983	0.9211	0.9322	0.7259	0.7724

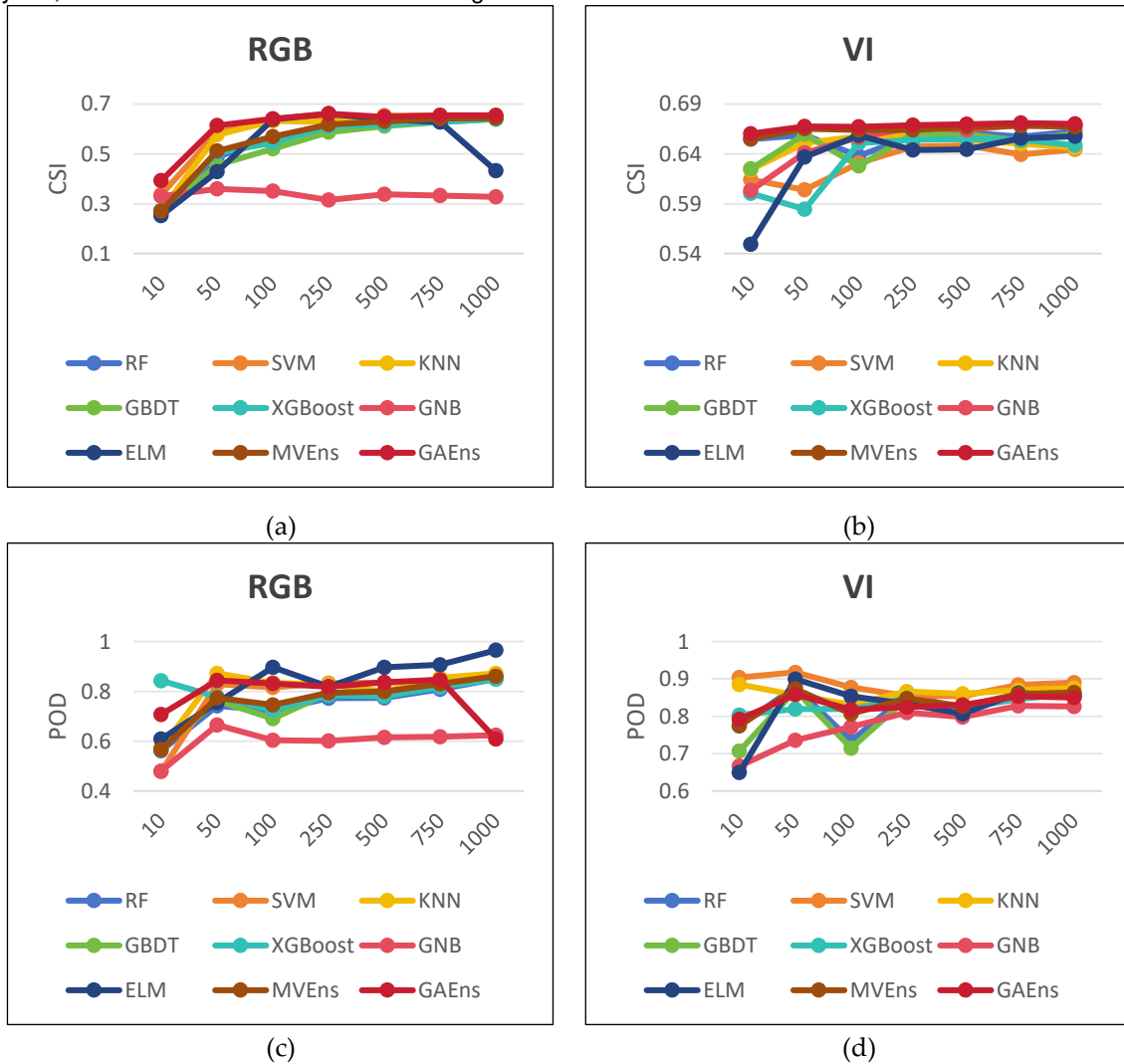
Table 3. Classification accuracy with morphological post-processing.

Discussion

The effect of sample size on classification results

In classification tasks, it is generally observed that a larger training sample size leads to higher final classification accuracy. However, the increase in samples also results in longer training times and higher manual costs for sample selection. In the context of monitoring pests and diseases in CCB trees as discussed in this paper, striking a balance between sample size and accuracy is crucial to enhance the algorithm's practicality. Therefore, the study sets the training samples at 10, 50, 100, 250, 500, 750, and 1000, evaluating the algorithm's accuracy performance across various sample sizes to determine the optimal number of samples.

By employing the aforementioned methodology, the accuracy of the sample sizes mentioned earlier was analyzed, and the outcomes are illustrated in Figure 5:



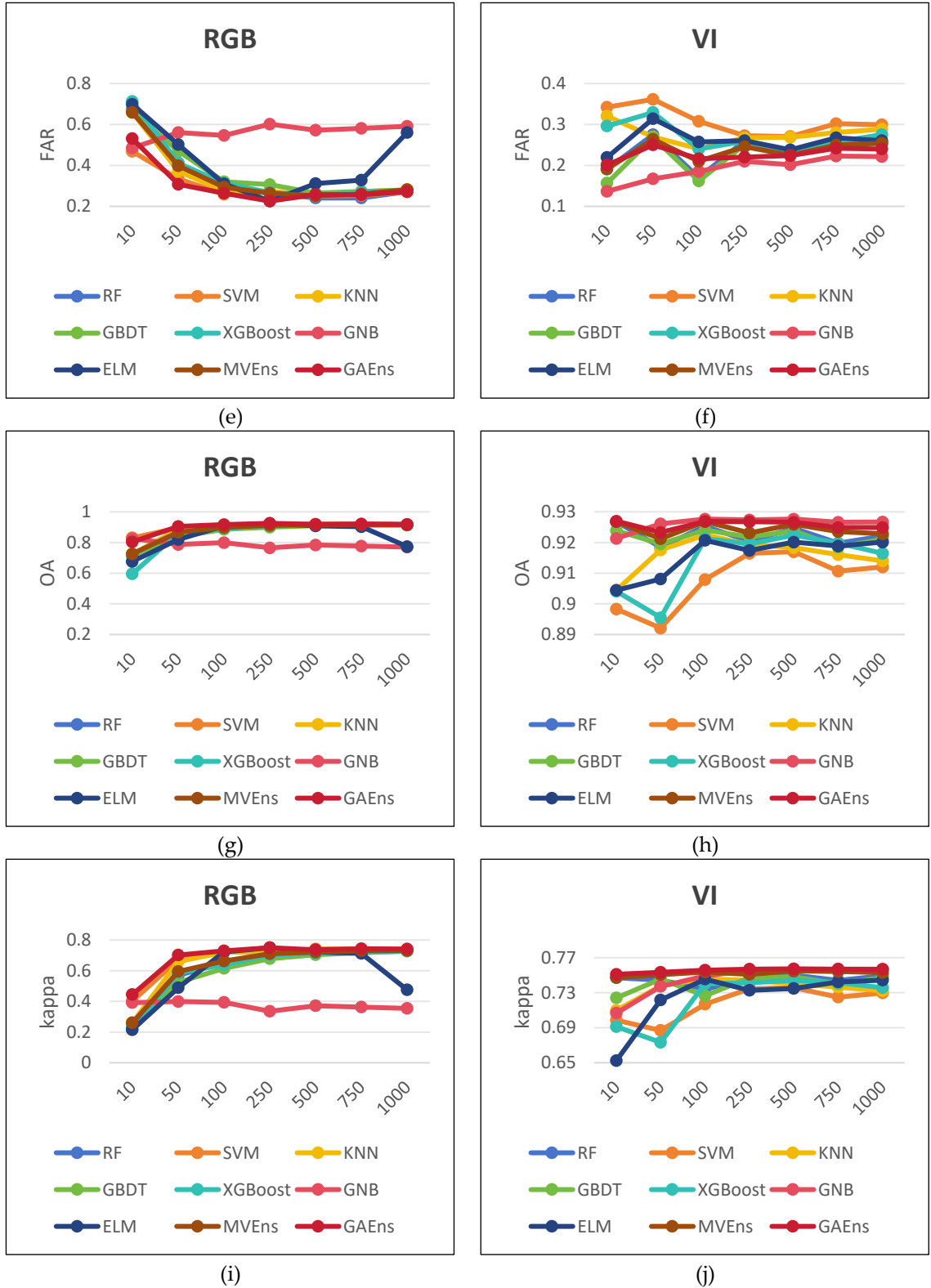


Figure 5. Objective evaluation results with different samples. (a) RGB-CSI; (b) VI-CSI; (c) RGB-POD; (d) VI-POD; (e) RGB-FAR; (f) VI-FAR; (g) RGB-OA; (h) VI-OA; (i) RGB-kappa; (j) VI-kappa.

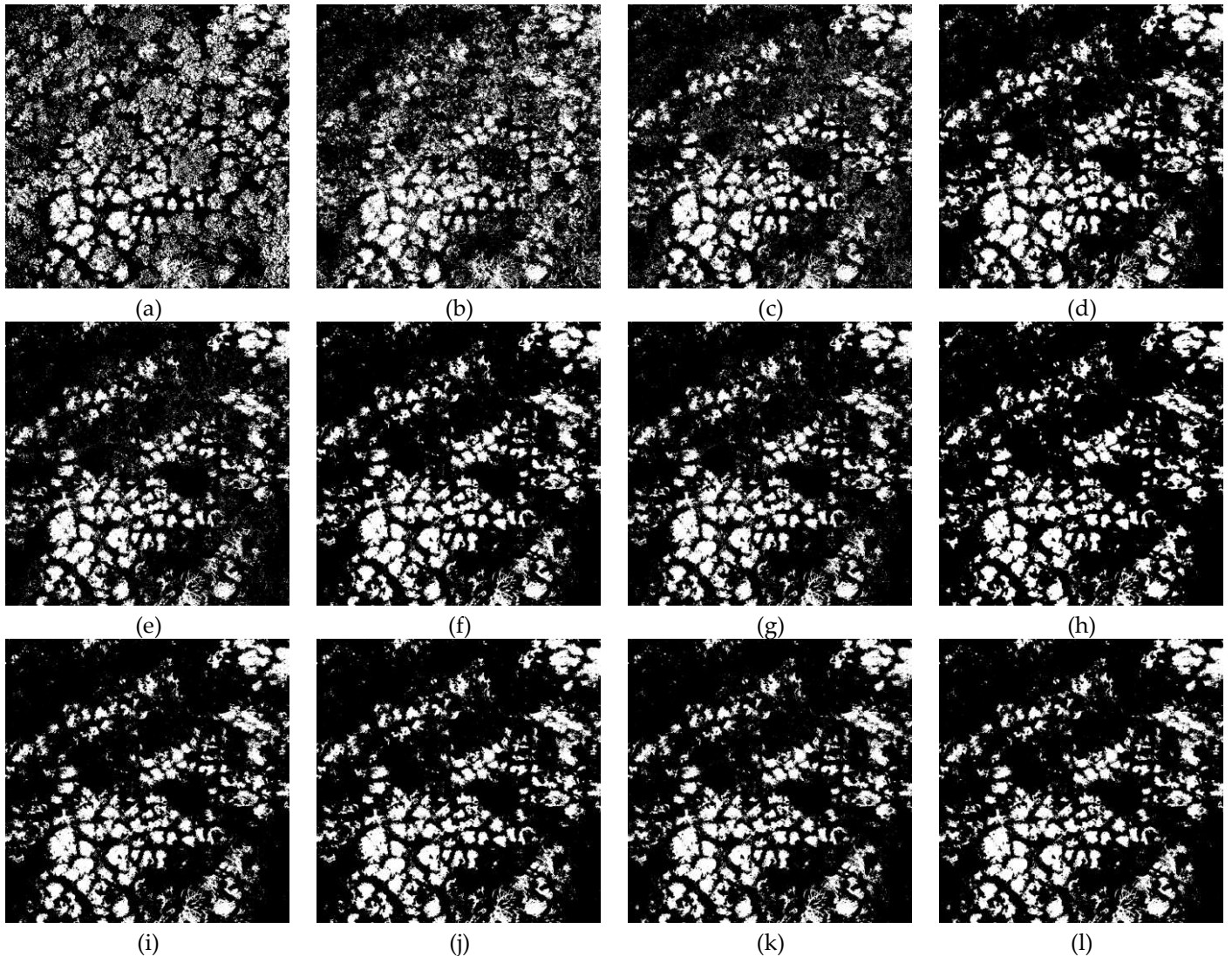
From Figure 5, it is evident that as the number of samples increases, the accuracy values of CSI, POD, OA, and kappa for both the original RGB and RGB-derived vegetation indices data sources exhibit improvement. Initially, the original RGB data source demonstrates a notable enhancement in accuracy with the growing sample size, followed by a plateau in the performance of the classification model where accuracy tends to stabilize. Conversely, the RGB-derived vegetation indices showcase high classification accuracy within the sample size range of 10-1000. While there is a significant initial improvement in accuracy with the increasing sample size, it is relatively smaller compared to the original RGB data source. Ultimately, both data sources

reach a saturation point in accuracy. Notably, in the early stages with fewer samples, the accuracy of the RGB-derived vegetation indices significantly surpasses that of the original RGB data source. Furthermore, the utilization of various machine learning algorithms under the vegetation indices leads to enhanced concentration and stability, resulting in improved classification outcomes for ensemble voting and overall accuracy. Regarding the FAR metric, both the original RGB data source and the vegetation indices data source demonstrate a decrease in classification error rates as the sample size increases, indicating that a higher number of samples leads to more precise results.

Upon analyzing the CSI, OA, and kappa indicators, it becomes apparent that with an increasing number of samples, the accuracy of direct majority voting ensembles does not consistently stabilize above that of individual machine learning algorithms. In contrast, voting strategies leveraging genetic algorithms consistently outperform most machine learning algorithms, showcasing superiority and stability in the line chart. From the accuracy trends across different sample sizes, it is evident that the accuracy enhancement beyond 50 samples is not substantial. Hence, for the RGB-derived vegetation indices data source, 50 samples serve as a favorable balance between accuracy and cost considerations.

The advantages of the genetic algorithm integration strategy

To enhance the accuracy of the final ensemble strategy, this study refines the voting approach by integrating genetic algorithms and explores seven distinct sample sizes - 10, 50, 100, 250, 500, 750, and 1000. The outcomes of this refined strategy are depicted in Figures 6 and 7, illustrating the impact across varying sample sizes.



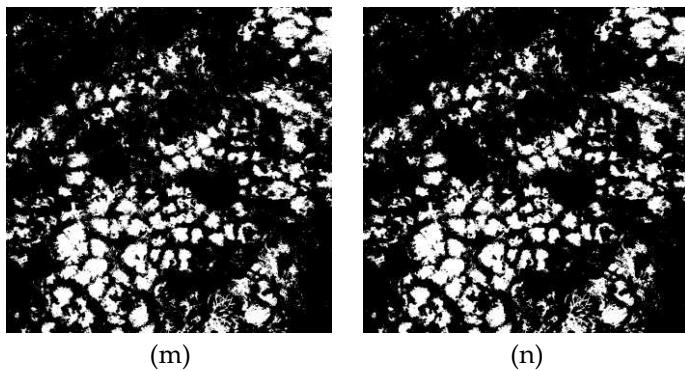
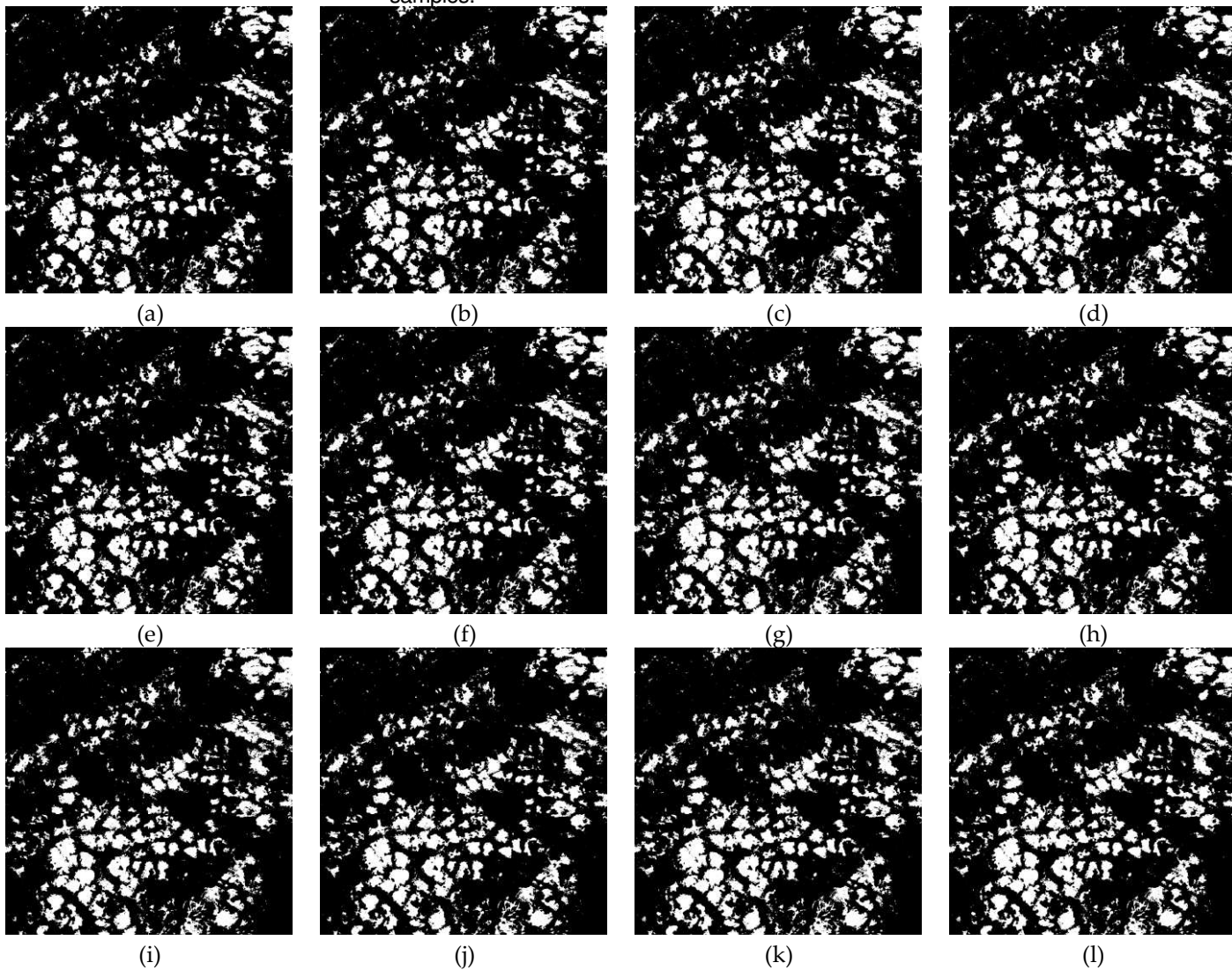


Figure 6. The ensemble strategy results for the original RGB image. (a) MVEns with 10 samples; (b) GAEns with 10 samples; (c) MVEns with 50 samples; (d) GAEns with 50 samples; (e) MVEns with 100 samples; (f) GAEns with 100 samples; (g) MVEns with 250 samples; (h) GAEns with 250 samples; (i) MVEns with 500 samples; (j) GAEns with 500 samples; (k) MVEns with 750 samples; (l) GAEns with 750 samples; (m) MVEns with 1000 samples; (n) GAEns with 1000 samples.



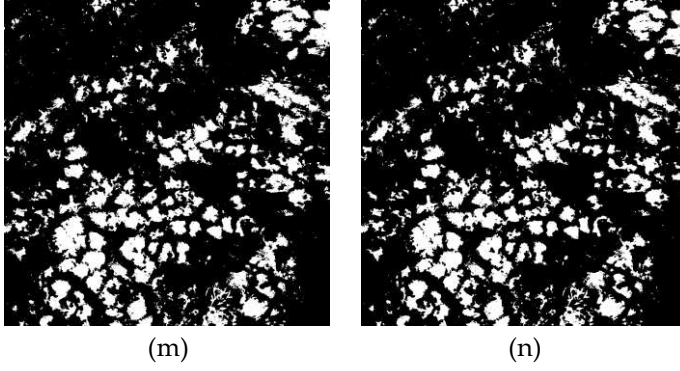
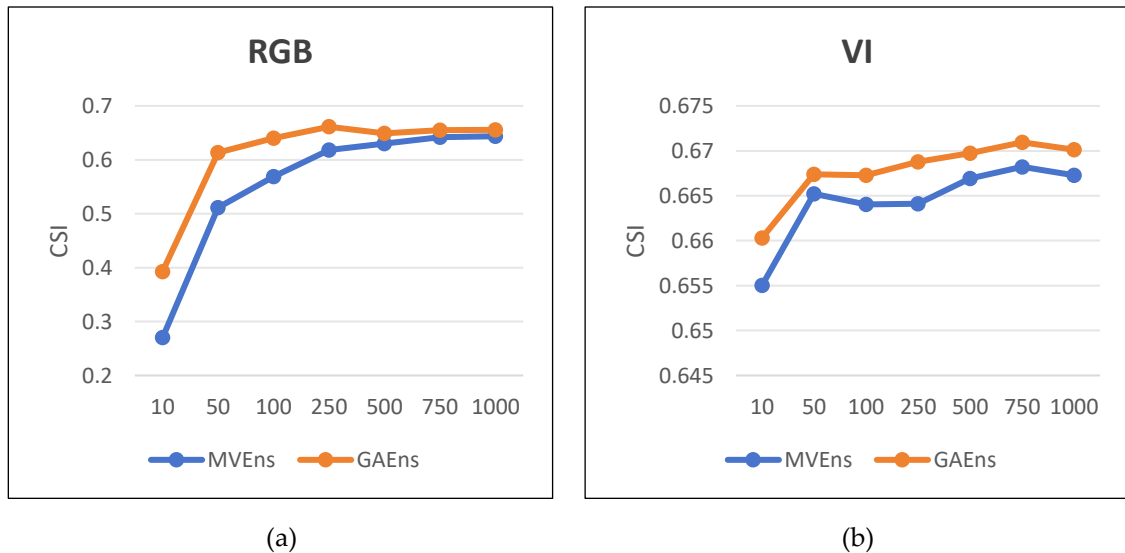


Figure 7. The ensemble strategy results for the RGB-derived vegetation indices. (a) MVEs with 10 samples; (b) GAEns with 10 samples; (c) MVEs with 50 samples; (d) GAEns with 50 samples; (e) MVEs with 100 samples; (f) GAEns with 100 samples; (g) MVEs with 250 samples; (h) GAEns with 250 samples; (i) MVEs with 500 samples; (j) GAEns with 500 samples; (k) MVEs with 750 samples; (l) GAEns with 750 samples; (m) MVEs with 1000 samples; (n) GAEns with 1000 samples.

Upon comparing the outcomes of the majority voting ensemble strategy with the genetic algorithm ensemble strategy, a noticeable difference emerges. The genetic algorithm ensemble strategy produces result images with smoother and more continuous edges in the white areas representing diseased trees, in contrast to the jagged edges present in the majority voting classification images. The genetic algorithm not only reintroduces valuable information that may have been overlooked by the majority voting approach but also rectifies misclassified regions, thereby refining the classification results and minimizing the risk of information loss or omission during the classification process. Furthermore, the efficiency of the genetic algorithm shows improvement as the sample size increases.

It is important to highlight that the genetic algorithm-based voting strategy and the direct majority voting strategy results discussed here do not involve morphology post-processing. Consequently, in cases where classification images exhibit significant noise, optimization can be accomplished through morphological filtering in subsequent stages. Additionally, it is worth noting that when dealing with original RGB data sources, noise levels are pronounced at a sample size of 10; however, the genetic algorithm voting strategy demonstrates the ability to mitigate noise to some extent. As the sample size increases, noise levels gradually diminish. Conversely, the two voting ensemble strategies based on RGB-derived vegetation indices exhibit considerably less noise, underscoring the superior capability of vegetation indices data sources in extracting pest and disease features effectively, thereby enhancing monitoring outcomes.

As illustrated in Figure 8, the accuracy variation results for all the aforementioned sample sizes are depicted.



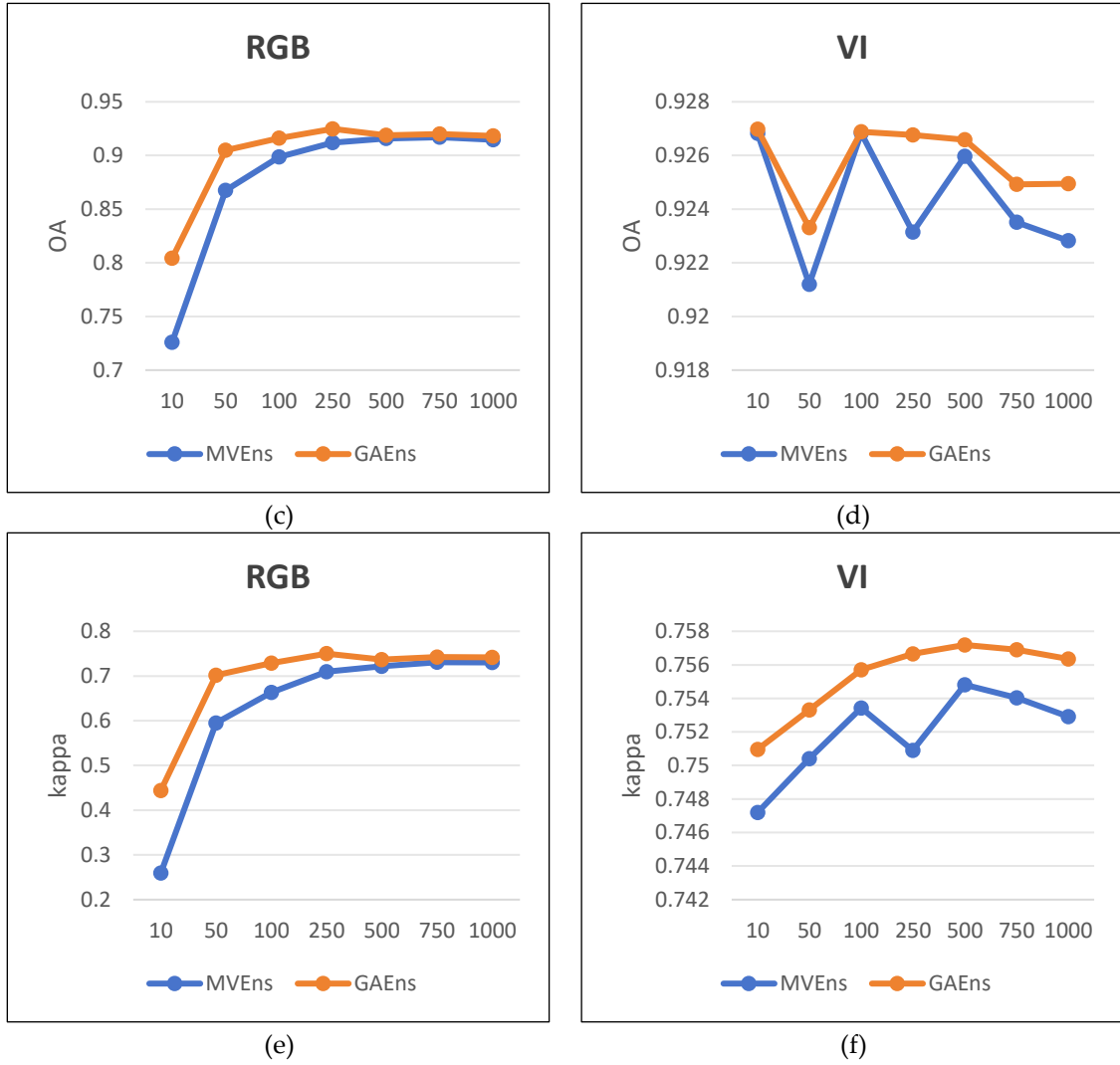


Figure 8. Accuracy variation results of the ensemble strategies. (a) RGB-CSI; (b) VI-CSI; (c) RGB-OA; (d) VI-OA; (e) RGB-kappa; (f) VI-kappa.

From the line chart, it is evident that the accuracy of the genetic algorithm voting strategy surpasses that of the direct majority voting strategy for both CSI, OA, and kappa indices across varying sample sizes. Specifically, when analyzing the original RGB data source, the genetic algorithm voting strategy initially holds a clear advantage as the sample size increases, but this advantage gradually diminishes and eventually stabilizes. On the other hand, when examining the RGB-derived vegetation indices data source, the direct majority voting strategy consistently achieves high accuracy across all sample sizes, yet the genetic algorithm voting strategy continues to outperform it significantly.

More application scenario data

In order to assess the algorithm's versatility and applicability, the genetic algorithm voting strategy model developed in section 3, utilizing RGB-derived vegetation indices data with a sample size of 50, was directly tested in new scenarios. Initially, a range of machine learning models including RF, SVM, KNN, GBDT, XGBoost, GNB, and ELM were employed to generate classification maps for the new scenarios. Subsequently, following morphological post-processing, the optimized weights derived from the genetic algorithm were utilized for ensemble voting to produce the final classification map. The outcomes are detailed in Figure 9. The results demonstrate the adaptability of the previously established model across diverse scenarios, although discrepancies between predicted outcomes and the ground truth map persist in certain intricate and uncertain regions.

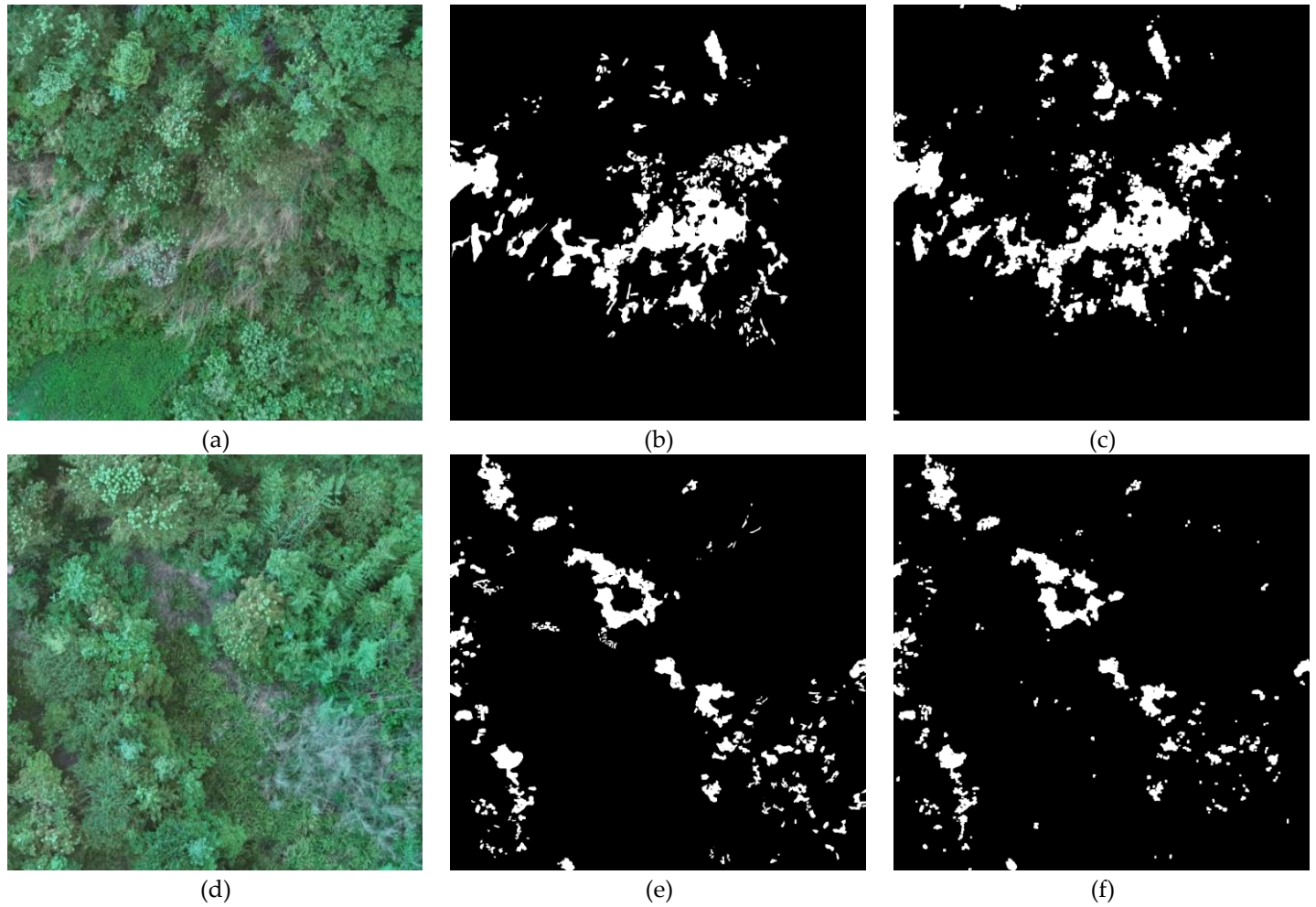


Figure 9. Results of other scenarios. (a) the original image of scene 1; (b) the ground truth of scene 1; (c) the result of scene 1; (d) the original image of scene 2; (e) the ground truth of scene 2; (f) the result of scene 2.

The accuracy indicators, including CSI, OA, and Kappa, for the two scenarios were computed using the confusion matrix, and the findings are presented in Table 4:

	CSI	OA	Kappa
Scene 1	0.6734	0.9635	0.7847
Scene 1	0.6695	0.9777	0.7903

Table 4. Classification accuracy of other scenario data.

Taking into account data from both scenarios, the model trained on the original dataset demonstrates high accuracy in diverse scenarios and can effectively pinpoint pest-infested areas in visible light images. Given that the OA indicator encompasses both pest and non-pest regions, and the non-pest regions have a notably larger extent, the OA accuracy metric yields a very high value. Nonetheless, on the whole, the model trained in the preceding section exhibits commendable performance in the new scenario, showcasing strong generalization capabilities.

Conclusion

Compared to multispectral/hyperspectral images, UAV RGB images are more cost-effective and easier to acquire, making them particularly well-suited for applications in large and challenging-to-access areas. The study is centered around a vast CCB plantation area, where the CCB trees grow tall and the mountainous terrain presents steep and arduous paths to navigate. Utilizing UAV RGB images for the classification of pest and disease trees is deemed more suitable in this context. The paper introduces RGB-derived vegetation indices as features for model training, combines various machine learning algorithms such as RF, SVM, KNN, GBDT, XGBoost, GNB, and ELM through ensemble techniques, and proposes the utilization of genetic

algorithm based ensemble strategy to optimize and assign weights to the classification results obtained from multiple machine learning algorithms. Through a series of experiments, it has been demonstrated that these derived vegetation indices offer more valuable information for detecting pests and diseases compared to directly using RGB bands. Even with a limited number of training samples, the approach yields stable and highly precise classification outcomes, with reduced noise and misclassification. While numerous vegetation indices can be derived from RGB images, and the monitoring effectiveness based on these indices surpasses that of the original RGB images, it is evident that the information extraction capacity of derived vegetation indices remains constrained by the limited information present in the RGB original images. Leveraging multispectral/hyperspectral images continues to be a superior strategy for enhancing monitoring capabilities, as these images offer a wider array of spectral information, facilitating more precise identification of pests and diseases across multiple spectral dimensions. Taking into account various cost considerations such as economic, usage, and computational costs, utilizing UAV RGB images for high-precision pest and disease monitoring represents a sound approach.

Methods

Study Area

Xinhuan Dong Autonomous County, formerly known as Huang County, is situated in Huaihua City, Hunan Province. Nestled in the western part of the Xiangxi low mountainous region, it borders the Yungui Plateau to the west and connects with Zhijiang to the east. The county shares its borders with Guizhou on the southwest, northwest, and north sides, covering a total area of 1,508 square kilometers, with mountainous terrain encompassing 84.22% of the region. The highest peak, Tianlei Mountain, stands at an elevation of 1,136 meters. Xinhuan County is distinguished by its numerous streams, creating a landscape of valleys and plains. The forest coverage in Xinhuan reaches 67.5%, making it the exclusive source of right-handed CCB trees in mainland China. Geographically, Xinhuan County spans from 108° 47'13" E to 109° 26'45" E longitude and 27° 4'16" N to 27° 29'58" N latitude, stretching 52.5 kilometers from east to west and 42.3 kilometers from north to south. The county experiences a subtropical monsoon humid climate, characterized by well-defined seasons, ample precipitation, and moderate humidity. The average annual temperature stands at 16.6° C, with an average annual precipitation of 1,160.7 millimeters and a total sunshine duration ranging from 1,014.5 to 1,590.2 hours. There are 297.4 frost-free days annually. Spring weather is variable, often marked by cold wave invasions. Summers and autumns tend to be dry, while the severe cold period is brief, and the frost-free period is extended. Xinhuan County's climate and geographical features make it conducive for agriculture, forestry, animal husbandry, and fishing. The region's favorable conditions, including suitable temperature, precipitation, and abundant mountain resources, create an optimal environment for the growth of CCB trees, ensuring an ample supply of raw materials for the production and advancement of Xinhuan CCB.

Data Acquisition

The UAV utilized in this study is the DJI Mavic 3M, depicted in Figure 10. This UAV serves as an aerial survey platform that combines a visible light camera and a multispectral camera, facilitating precise agricultural production management. The DJI Mavic 3M features a 20-megapixel visible light camera and four 5-megapixel multispectral cameras (green, red, red edge, and near-infrared) integrated into its design. Additionally, the UAV incorporates a multispectral light intensity sensor for capturing solar irradiance. During the 2D reconstruction process, the image data can be adjusted for lighting conditions to enhance the accuracy of NDVI results, ensuring improved precision and consistency in data collection across various timeframes. Equipped with an RTK module, the DJI Mavic 3M achieves centimeter-level high-precision positioning, enabling image-free control aerial surveying even during high-speed flights. The UAV has a maximum takeoff weight of 1050 grams and dimensions measuring 347.5mm * 283mm * 139.6mm. It boasts a maximum hover time of 37 minutes and can achieve a maximum horizontal flight speed of 15 meters per second.



Figure 10. DJI Mavic 3M UAV Image.

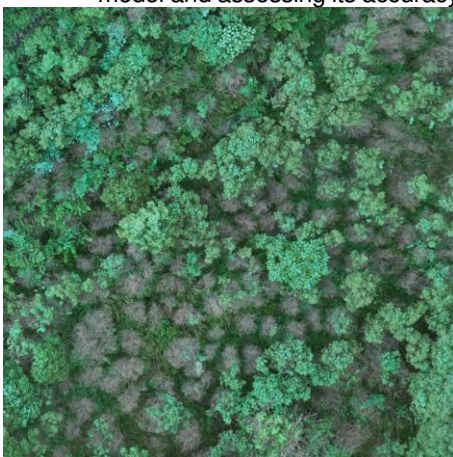
The data was gathered at the CCB tree planting base in Xinhuan County on July 13th, 2023, at 3:00 PM. The weather conditions were favorable with clear skies, and the UAV flew at an altitude of 80 meters. Both the forward overlap and side overlap rates were set at 80% and 70% respectively to ensure comprehensive coverage and data

accuracy. The photography mode employed was "equal time interval photography" to enhance image overlap and maintain data integrity throughout the survey. Following processing using the DJI Zhitu software, the RGB data acquired generated a composite image of the entire area, displayed in Figure 11. The shadowed region with darker colors in the image signifies the mountain's rear section.



Figure 11. The overall RGB image of CCB tree planting base in Xinhuan County.

Due to the extensive size of the area under study, three smaller regions have been chosen as focal points for the algorithm, depicted in Figure 12. Notably, the images reveal the presence of visibly deteriorated plants, likely impacted by pests and diseases. To establish ground truth data, the precise locations of the affected plants have been manually delineated, assigning a grayscale value of 255 to the pest and disease-affected areas, while the unaffected normal plant regions are designated as 0. These ground truth images serve as the basis for training the model and assessing its accuracy.



(a)



(b)



(c)

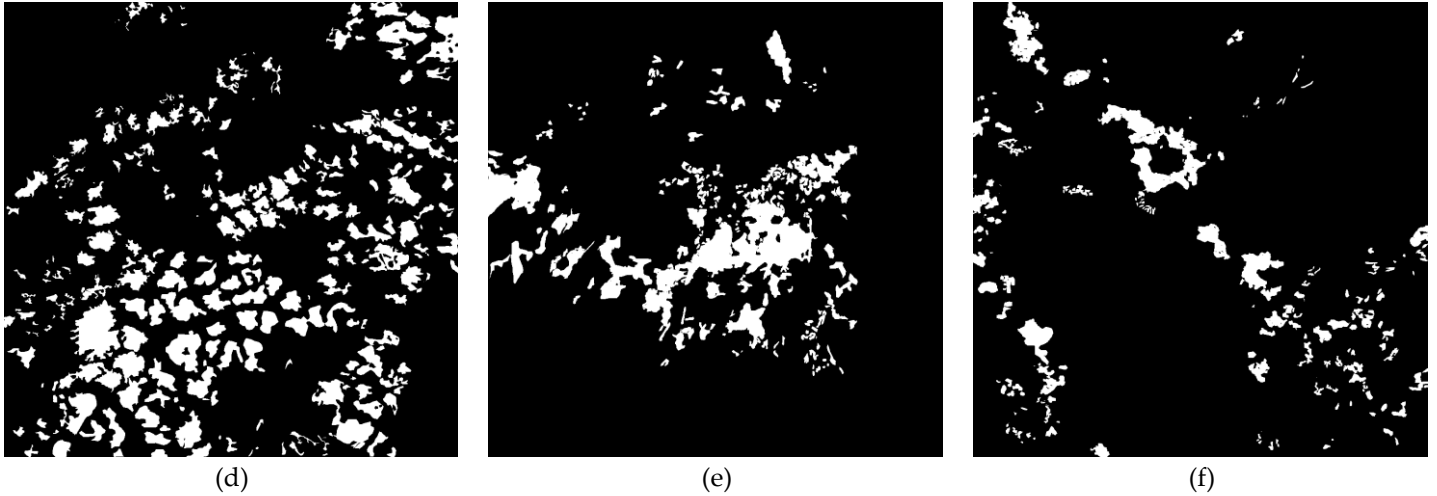


Figure 12. The original RGB images and the ground truth. (a) Scene 1; (b) Scene 2; (c) Scene 3; (d) Ground truth image for Scene 1; (e) Ground truth image for scene 2; (f) Ground truth image for scene 3.

Method

The comprehensive flowchart of this paper is illustrated in Figure 13 below, encompassing key steps such as vegetation index computation, machine learning algorithm classification, post-classification processing, and the optimization of ensemble strategies through genetic algorithms.

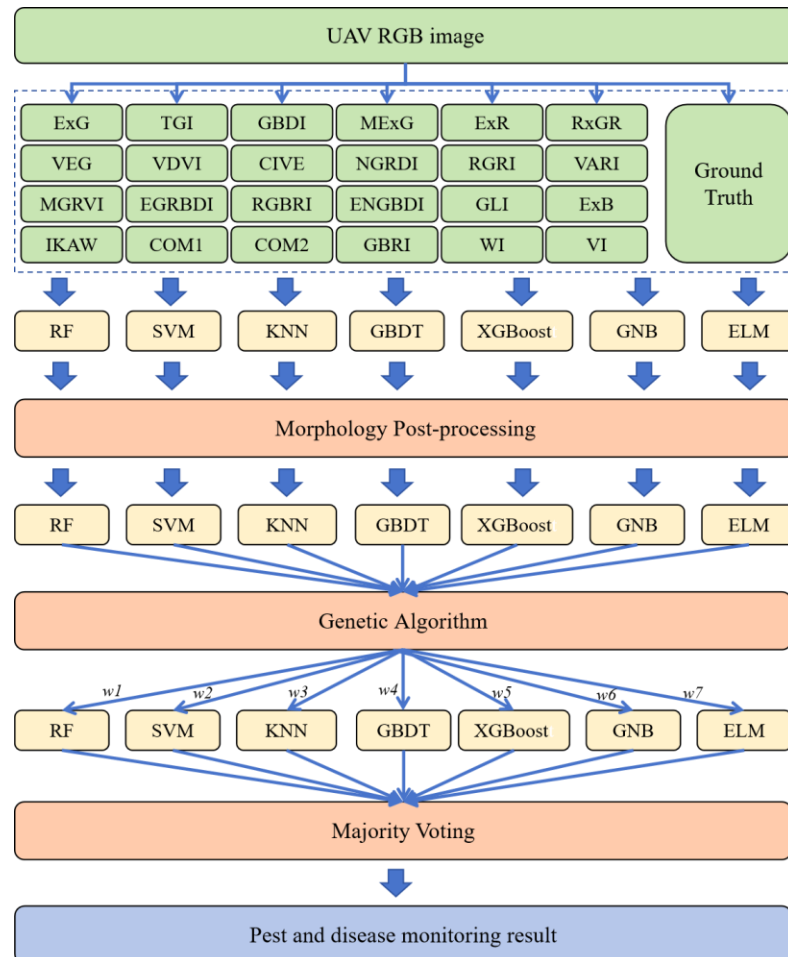


Figure 13. Overall flowchart.

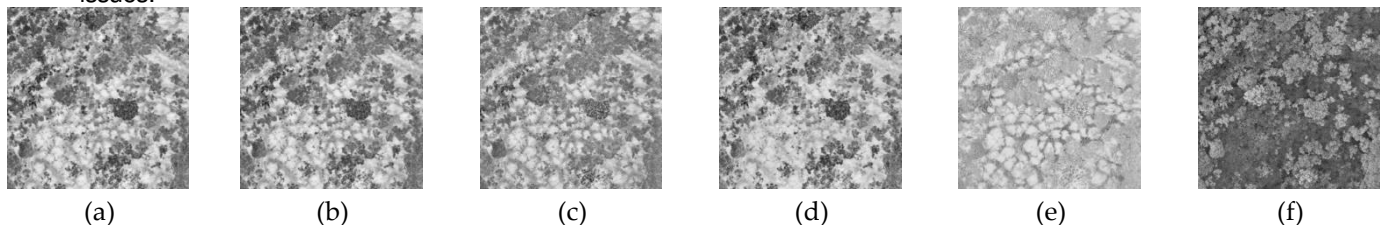
RGB-derived Vegetation Indices

This paper adopts 24 commonly utilized vegetation indices derived from RGB images, as suggested in existing literature, along with the following formulaic representations shown in table 5:

Name	Expression	Reference
Excess green index(EXG)	$EXG = 2 * \text{green} - \text{red} - \text{blue}$	[22]
Triangular greenness index(TGI)	$TGI = \text{green} - 0.39 * \text{red} - 0.61 * \text{blue}$	[23]
Green-blue difference index(GBDI)	$GBDI = \text{green} - \text{blue}$	[24]
Modified excess green index(MExG)	$MExG = 1.262 * \text{green} - 0.844 * \text{red} - 0.311 * \text{blue}$	[25]
Excess green index(ExR)	$ExR = 1.4 * \text{red} - \text{green}$	[26]
Excess Green minus Excess Red(ExGR)	$ExGR = ExG - ExR$	[27]
Vegetativen(VEG)	$VEG = \text{green} / (\text{red}^a * \text{blue}^{1-a})$ $a=0.667$	[28]
Visible-band difference vegetation index(VDVI)	$VDVI = (2 * \text{green} - \text{red} - \text{blue}) / (2 * \text{green} + \text{red} + \text{blue})$	[29]
Color Index of Vegetation Extraction(CIVE)	$CIVE = 0.441 * r - 0.811 * g + 0.385 * b + 18.78745$	[30]
Normalized Green-red Difference Index(NGRDI)	$NGRDI = (\text{green} - \text{red}) / (\text{green} + \text{red})$	[31]
Red Green Ratio Index(RGRI)	$RGRI = \text{red} / \text{green}$	[32]
Visible Atmospherically Resistant Index(VARI)	$VARI = (\text{green} - \text{red}) / (\text{green} + \text{red} - \text{blue})$	[33]
Modify Green-Red Vegetation Index(MGRVI)	$MGRVI = (\text{green}^2 - \text{red}^2) / (\text{green}^2 + \text{red}^2)$	[34]
Excess green-red-blue difference index (EGRBDI)	$EGRBDI = (4 * \text{green}^2 - \text{red} * \text{blue}) / (4 * \text{green}^2 + \text{red} * \text{blue})$	[35]
Red-Green-Blue Ratio Index (RGBRI)	$RGBRI = (\text{green}^2 - \text{red} * \text{blue}) / (\text{green}^2 + \text{red} * \text{blue})$	[36]
Enhance Normalized Red-Blue Difference Index(E-NGBDI)	$E-NGBDI = (\text{green}^2 - \text{blue}^2) / (\text{green}^2 + \text{blue}^2)$	[37]
Green leaf index (GLI)	$GLI = (2 * g - r - b) / (2 * g + r + b)$	[38]
Excess blue index(ExB)	$ExB = (1.4 * \text{blue} - \text{green}) / (\text{green} + \text{red} + \text{blue})$	[39]
kawashima index (IKAW)	$IKAW = (\text{red} - \text{blue}) / (\text{red} + \text{blue})$	[40]
Comprehensive 1(COM1)	$COM1 = ExG + CIVE + ExGR + VEG$	[41]
Comprehensive 1(COM2)	$COM2 = 0.36 * ExG + 0.47 * CIVE + 0.17 * VEG$	[42]
Green-Blue Ratio Index(GBRI)	$GBRI = \text{blue} / \text{green}$	[43]
Woebbecke index(WI)	$WI = (\text{green} - \text{blue}) / (\text{red} - \text{green})$	[44]
VI	$VI = (2 * g - r - b) - (1.4 * r - g)$	[45]

Table 5. The 24 vegetation indices.

Each vegetation index captures the unique characteristics of vegetation under specific conditions. In the case of 24 vegetation indices derived from RGB bands, each index was individually standardized to generate grayscale images, as illustrated in Figure 14. The analysis reveals that certain vegetation indices, such as GLI, RGRI, and EGRBDI, exhibit a heightened sensitivity to vegetation impacted by pests and diseases. Conversely, ExGR demonstrates a greater responsiveness to healthy plants, whereas indices like IKAW and WI may not offer as distinct responses to vegetation conditions. By amalgamating the sensitivities of multiple vegetation indices to plants affected by pests and diseases, it becomes feasible to achieve precise monitoring and detection of such issues.



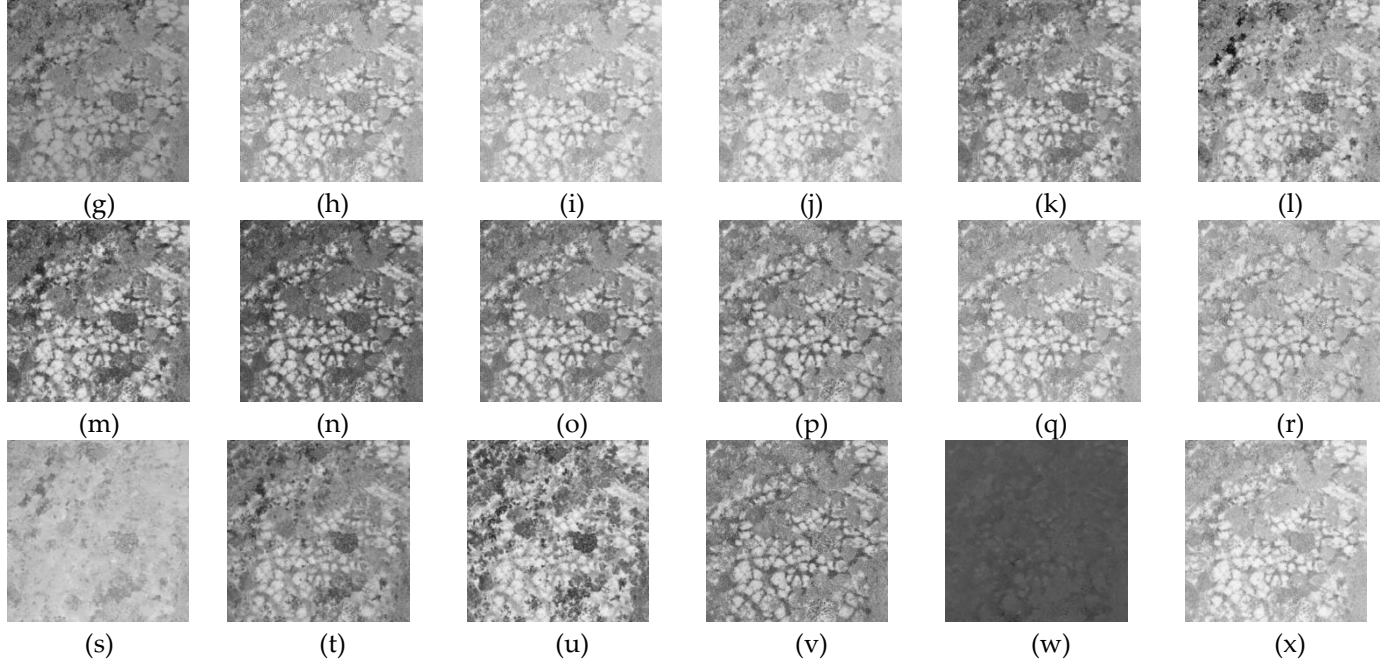


Figure 14. Gray-scale images of 24 vegetation indices. (a) ExG; (b) TGI; (c) GBDI; (d) MExG; (e) ExR; (f) ExGR; (g) VEG; (h) VEG; (i) CIVE; (j) NGRDI; (k) RGRDI; (l) VARI; (m) MGRVI; (n) EGRBDI; (o) RGBRI; (p) E-NGBDI; (q) GLI; (r) ExB; (s) IKAW; (t) COM1; (u) COM2; (v) GBRI; (w) WI; (x) VI index.

Classification

This study leverages a diverse array of machine learning algorithms, including RF, SVM, KNN, GBDT, XGBoost, GNB, and ELM, to classify diseased and healthy plants and subsequently conduct comprehensive image-based pest and disease detection. Subsequently, an ensemble approach, optimized through a genetic algorithm, is employed to enhance the final monitoring of pest and disease outbreaks. Prior to model training, the creation of training samples is imperative. By utilizing the 24 vegetation indices as features, with diseased and healthy plants designated as positive and negative labels, a specific number of pixels (e.g., 50, 100, etc.) are randomly selected to form the training dataset. Consequently, each training sample comprises the following components:

$$\langle x_1, x_2, \dots, x_{24}, 0 \rangle \text{ or } \langle x_1, x_2, \dots, x_{24}, 1 \rangle, \quad (6)$$

where each of the 24 vegetation indices is denoted as unnormalized, with the label "0" assigned to healthy plants and the label "1" designated for diseased plants.

Below are the 7 machine learning algorithms employed in this paper:

. Random Forest (RF)

RF [46] is a classifier based on ensemble learning. It integrates multiple weak decision trees together to obtain the final result through voting, achieving higher accuracy and generalization. RF is formed by combining multiple decision trees $\{h(x, \theta_t), t = 1, 2, \dots, T\}$. Where θ_t is a random variable that follows an independent distribution, x is the independent variable, and T is the number of decision tree nodes. The prediction result of the classification model is:

$$\bar{H}(x) = \arg \max_{Y} \sum_{i=1}^T I(h(x, \theta_i) = Y), \quad (7)$$

where $\bar{H}(x)$ is the prediction result, Y is the classification type, I is the indicator function.

. SVM

SVM [47] is a binary classification model. Its fundamental model is a linear classifier with the maximum margin defined in the feature space, setting it apart from the perceptron. SVM also incorporates kernel tricks, transforming it into a non-linear classifier. The learning approach of SVM aims to maximize the margin, a process that can be mathematically expressed as solving a convex quadratic programming problem or, equivalently, minimizing a

regularized hinge loss function. The learning algorithm of SVM is an optimization technique designed for solving convex quadratic programming problems.

. KNN

The K-Nearest Neighbors algorithm, commonly known as KNN, is often described as a non-parametric lazy training algorithm. It operates without making any assumptions about the data distribution, such as adhering to a Gaussian distribution or being linearly separable, hence its non-parametric nature. This flexibility is advantageous as real-world data often deviate from traditional theoretical assumptions. Moreover, KNN does not necessitate training data for generalization; during the training phase, it solely involves storing training data, with computations deferred until the testing phase [48]. For test set data, KNN first calculates the distance between each sample point in the test set and the training set, employing a designated distance measurement formula for statistical evaluation. Subsequently, the samples are arranged based on the computed distances. The K samples with the smallest distances from the training set are chosen, and their labels are tallied and organized, with the most frequent label determining the final classification category. Typically, the value of K is set to be less than 20, and when K equals 1, classification is solely based on the nearest point.

. Gradient Boosting Decision Trees (GBDT)

GBDT is a forest-like structure composed of decision trees that leverages gradient boosting for learning and prediction. In GBDT, predictions are generated by aggregating the outputs of all the individual subtrees, as it operates as an ensemble model. The forest in GBDT is built sequentially, with decision subtrees generated one by one. Each new subtree is created by considering the residuals between the actual sample label values and the predicted values from the current forest.

. XGBoost

XGBoost [49] is a decision tree-based algorithm, akin to Random Forest (RF). Nevertheless, it sets itself apart from RF by iteratively improving the model's predictive performance during training through loss function minimization. In Python, the implementation of this algorithm relies on the scikit-learn 1.3.1 package. Key parameters to tune include the maximum depth of individual trees (`max_depth`), the fraction of features considered for node splits (`colsample_bytree`), the number of trees in the model (`n_estimators`), the fraction of training data used in each boosting iteration (`subsample`), and the learning rate (`learning_rate`). These parameters can be fine-tuned using a validation set.

. Gaussian Naive Bayes (GNB)

GNB is a machine learning classification technique that leverages probabilistic methods and operates under the assumption of Gaussian distribution. In GNB, it is assumed that each parameter, often referred to as a feature or predictor variable, independently influences the predictive capacity of the output variable. The prediction generated is a collective result of predictions made by all parameters, offering the probability of the dependent variable falling into each category. The ultimate classification is assigned to the group or class with the highest probability.

. Extreme Learning Machine (ELM)

ELM is a training method designed for Single Hidden Layer Feedforward Neural Networks (SLFN). Unlike traditional algorithms used for training SLFNs, ELM employs a unique approach by randomly selecting the input layer weights and hidden layer biases. The output layer weights, on the other hand, are determined analytically by minimizing a loss function that encompasses the training error term along with a regularization term based on the norm of the output layer weights. This computation of the output layer weights is based on the principles of the Moore-Penrose (MP) generalized inverse matrix theory. Research has shown that ELM retains the universal approximation capabilities of SLFNs, even when incorporating randomly generated hidden layer nodes. Over the last decade, ELM has attracted significant interest in both theoretical studies and practical applications. In terms of learning efficiency, ELM offers numerous benefits, such as a decreased number of training parameters, enhanced learning speed, and robust generalization capabilities.

. Ensemble Learning

The process of constructing ensemble learning models typically involves two main steps: generating base predictors and selecting an appropriate ensemble strategy. The cornerstone of creating a reliable ensemble model lies in the careful selection of base predictors. Common ensemble strategies encompass methods such as voting, bagging, and boosting. A voting classifier is a type of ensemble learning algorithm that consolidates the predictions of multiple classifiers, each of which has been trained successfully. It is crucial that the base classifiers differ from each other, as employing diverse algorithms for the base classifiers enables the ensemble model to learn more about their individual errors, consequently enhancing overall accuracy. There exist two categories of voting classifiers: hard voting classifiers and soft voting classifiers. In hard voting, each base classifier casts a vote for their predicted result, and the class with the highest number of votes is chosen as the final output. On the other hand, soft voting involves calculating the weighted average of the class probabilities predicted by each base classifier.

The class with the highest average probability is then selected as the output result. In practice, soft voting is generally considered more effective than hard voting and is applicable to a wider range of scenarios based on empirical evidence and experience.

Classification post-processing

Our research has revealed that when the sample size is limited, various classification methods tend to produce results with noticeable noise. As a result, it becomes imperative to apply post-processing techniques to mitigate the noise to a certain extent. In this study, we have employed the opening operation from mathematical morphology to enhance the quality and accuracy of the classification images.

In morphological operations, there are two basic operations: expansion and erosion. Assuming the original image is f and the structural operator is S , expansion operation $f \oplus S$ and erosion operation $f \ominus S$ are opposite operations, which can be represented by the following expressions, respectively:

$$(f \oplus S)(x, y) = \max_{p, q} [f(x - p, y - q) + S(p, q)] \quad (8)$$

$$(f \ominus S)(x, y) = \min_{p, q} [f(x + p, y + q) - S(p, q)] \quad (9)$$

The dilation operation enlarges the contours of the image, effectively expanding the target area within the processed image. Conversely, the erosion operation reduces the size of intricate details, leading to a more compact target area within the processed image.

Building upon the principles of expansion and erosion, morphological operations combine these two processes to create an opening operation, which entails an initial erosion followed by an expansion. The operation proceeds as follows:

$$(f \circ S) = (f \ominus S) \oplus S \quad (10)$$

The opening operation is capable of removing noise pixels in binary images, smoothing out the boundaries of regions, and ultimately yielding more precise boundaries of plants impacted by pests and diseases.

Ensemble strategy based on genetic algorithm

The concept of genetic algorithms was initially introduced by Holland [50] in the 1960s and 1970s, inspired by the evolutionary theory's survival of the fittest principle. Schaffer [51] introduced the first genetic algorithm, Vector Evaluator, in 1985. In 1992, Hajela [52] integrated weights into genetic algorithms, leading to the proposal of weight-based genetic algorithms. In recent years, genetic algorithms have found extensive applications in addressing machine learning challenges such as clustering [53], cluster ensembles [54], feature selection [55], parameter optimization [56], sample selection and reconstruction [57], ensemble strategies [58], and classification [59].

The traditional ensemble strategy typically relies on majority voting to determine the final classification outcome, assuming uniform classification performance and equal weights for each base classifier. In this study, seven machine learning algorithms, including random forest, were employed to generate individual results for pest and disease detection. Following the conventional majority voting approach, if four or more classifiers identify a pixel as belonging to a diseased or pest-infested plant, the final decision is made accordingly. However, this study posits that the classification abilities of each algorithm vary. By assigning distinct weights to each algorithm and employing majority voting, enhanced detection accuracy can be attained. Consequently, this paper introduces an ensemble strategy that leverages genetic algorithm optimization.

The primary procedure is shown in figure 15 and unfolds as follows:

Data input

Classification maps generated by seven machine learning algorithms (RF, SVM, KNN, GBDT, XGBoost, GNB, ELM) along with 100 positive and negative validation samples randomly selected, distinct from the training samples.

Initial population creation: Seven machine learning algorithms are associated with 7 weight parameters, each represented using a 6-bit binary code ranging from 0 to 63. Consequently, each chromosome is composed of 42 binary bits, with each 6-bit segment denoting a weight parameter. The process begins by randomly generating 100 chromosomes, with the binary encoding of each chromosome also being randomly generated. These 100 chromosomes collectively constitute the initial population.

Fitness calculation

To calculate the fitness of each chromosome in the current population, the following steps are taken. Firstly, decode each chromosome into 7 decimal weight values, where each 6-bit segment corresponds to a weight for one of the 7 machine learning algorithms. Next, for every pixel in the validation set, gather the classification results from the classification maps generated by the 7 machine learning algorithms. These results are then weighted using the respective weights assigned to each model. For instance, if the weights for the first and second models are 15 and 35 respectively, the classification results of these models for the pixel will be repeated 15 and 35 times respectively. Subsequently, the combined classification results are subjected to a majority vote strategy to determine the final classification of the pixel, where the class with the highest votes is chosen as the pixel's classification. After applying the majority vote to all validation pixels, the fitness function is evaluated using the kappa coefficient. This process assigns a fitness value to each chromosome based on its performance.

Selection

The selection process involves calculating the fitness values of all chromosomes in the population. Subsequently, the M individuals with the highest fitness values are chosen to advance to the next stage of operations. This selection mechanism ensures that the outstanding traits of the parents are maintained and passed on to the next generation.

Crossover

After the selection of the top M chromosomes, pairs of chromosomes are randomly chosen for the crossover operation. Initially, a crossover point is randomly determined. Subsequently, the chromosome segments before and after this point are exchanged between the two chromosomes, resulting in the creation of two new chromosomes. This process is iterated multiple times to generate an adequate number of new chromosomes to match the population size.

Mutation

Following the defined mutation probability, each chromosome within the population is chosen, and a specific number of binary digits are randomly flipped at selected positions. This process involves changing 0s to 1s or 1s to 0s, resulting in the creation of the mutated chromosome. Once all chromosomes have undergone mutation, a new population is established.

Fitness Recomputation

In this step, the fitness of every chromosome in the updated population is computed, and the chromosome with the highest fitness is retained. If a chromosome's fitness surpasses that of the current best chromosome, it replaces the existing best chromosome. This iterative process continues until the designated number of evolution iterations is achieved. Upon completion of the algorithm, the best chromosome is output, allowing for the extraction of optimal weights through decoding.

Output

The optimal weights are utilized to adjust the classification maps generated by the seven machine learning algorithms. Subsequently, the final classification results are produced by applying the principle of majority voting, as outlined in the fitness calculation step.

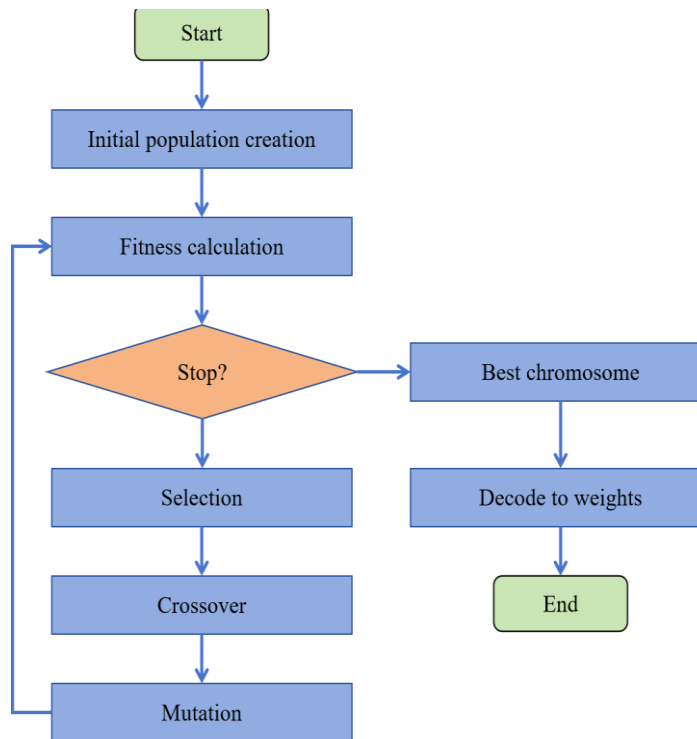


Figure 15. Flowchart of Genetic Algorithm.

Data availability

The data supporting the conclusions of this study are available from the corresponding author upon a reasonable request.

References

- Chen, L.; Su, J.; Li, L.; Li, B.; Li, W. A new source of natural D-borneol and its characteristic. *J. Med. Plants Res.* **5**(15), 7, (2011).
- Sun, X. M.; Liao, Q. F.; Zhou, Y. T.; Deng, X. J.; Xie, Z. Y. Simultaneous determination of borneol and its metabolite in rat plasma by GC-MS and its application to pharmacokinetic study. *J. Pharm. Anal.* **4**(05), 345-350, (2014).
- Chen, X. Y.; Zhao, X. N.; Zeng, H. F.; Xie, J. H.; Chen, X. L.; Liang, Y. Z.; Wu, Q.D.; Su, Z.R.; Wang, H.F.; Lai, X.P. Natural borneol recycling from Cinnamomum camphor chvar. Borneol oil residue by fractional distillation and recrystallization. *Trop. J. Pharm. Res.* **13**(9), 1463-1470, (2014).
- Zhang, T.; Zhang, X.; Liu, H.; Pei, X. Application of remote sensing technology in monitoring forest diseases and pests. *Plant Diseases and Pests.* **1**(3), 57-62, (2010).
- De Beurs, K. M.; Townsend, P. A. Estimating the effect of gypsy moth defoliation using MODIS. *Remote Sens. Environ.* **112**(10), 3983-3990, (2008).
- Xu, Z.; Zhang, Q.; Xiang, S.; Li, Y.; Huang, X.; Zhang, Y.; Guo, X. Monitoring the severity of Pantana phyllostachysae Chao infestation in Moso bamboo forests based on UAV multi-spectral remote sensing feature selection. *Forests.* **13**(3), 418, (2022).
- Jiang, J. B.; Chen, Y. H.; Huang, W. J. Using the distance between hyperspectral red edge position and yellow edge position to identify wheat yellow rust disease. *Spectrosc. Spect. Anal.* **30**(6), 1614-1618, (2010).
- Bendig, J.; Bolten, A.; Bennertz, S.; Broscheit, J.; Eichfuss, S.; Bareth, G. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. *Remote sens.* **6**(11), 10395-10412, (2014).
- Kazmi, W.; Garcia-Ruiz, F. J.; Nielsen, J.; Rasmussen, J.; Andersen, H. J. Detecting creeping thistle in sugar beet fields using vegetation indices. *Comput. and Electron. Agric.* **112**, 10-19, (2015).

10. Liu, T.; Li, R.; Jin, X.; Ding, J.; Zhu, X.; Sun, C.; Guo, W. Evaluation of seed emergence uniformity of mechanically sown wheat with UAV RGB imagery. *Remote Sens.* **9**(12), 1241, (2017).
11. Zhang, D.; Mansaray, L. R.; Jin, H.; Sun, H.; Kuang, Z.; Huang, J. A universal estimation model of fractional vegetation cover for different crops based on time series digital photographs. *Comput. Electron. Agric.* **151**(6), 93-103, (2018).
12. Görlich, F.; Marks, E.; Mahlein, A. K.; König, K.; Lottes, P.; Stachniss, C. UAV-based classification of cercospora leaf spot using RGB images. *Drones.* **5**(2), 34, (2021).
13. Zermas, D.; Nelson, H. J.; Stanitsas, P.; Morellas, V.; Mulla, D. J.; Papanikolopoulos, N. A methodology for the detection of nitrogen deficiency in corn fields using high-resolution RGB imagery. *IEEE Trans. on Autom. Sci. Eng.* **18**(4), 1879-1891, (2020).
14. Fawkes, C. A.; Leach, C. M. J.; Mathias, S.; Moore, A. P. Development of a data collection tool to profile osteopathic practice: Use of a nominal group technique to enhance clinician involvement. *Manual Therapy* **19**(2), 119-124, (2014).
15. Yue, Y.; Cheng, X.; Zhang, D.; Wu, Y.; Zhao, Y.; Chen, Y.; Fan, G.; Zhang, Y. Deep recursive super resolution network with Laplacian Pyramid for better agricultural pest surveillance and detection. *Comput. Electron. Agric.* **150**, 26-32, (2018).
16. Pan, C.; Xiao, D.; Lin, T.; Wang C. Classification and recognition for major vegetable pests in Southern China using SVM and region growing algorithm. *Transactions of the Chinese Society of Agricultural Engineering* **Vol.34(No.8)**, 192-199, (2018).
17. Cheng, X.; Zhang, Y.; Wu, Y.; Yue, Y. Agricultural Pests Tracking and Identification in Video Surveillance Based on Deep Learning. *Intelligent Computing Methodologies.* 58-70, (2017).
18. Mohanty, S. P.; Hughes, D. P.; Salathé, M. Using deep learning for image-based plant disease detection. *Front. plant sci.* **7**, 1419, (2016).
19. Healey, S. P.; Cohen, W. B.; Yang, Z.; Brewer, C. K.; Brooks, E. B.; Gorelick, N.; Hernandez, A. J.; Huang, C.; Hughes, M. J.; Kennedy, R. E.; Loveland, T. R.; Moisen, G. G.; Schroeder, T. A.; Stehman, S. V.; Vogelmann, J. E.; Woodcock, C. E.; Yang, L.; Zhu, Z. Mapping forest change using stacked generalization: An ensemble approach. *Remote Sens. Environ.* **204**, 717-728, (2018).
20. Xiao, Y.; Guo, Y.; Yin, G.; Zhang, X.; Shi, Y.; Hao, F.; Fu, Y. UAV multispectral image-based urban river water quality monitoring using stacked ensemble machine learning algorithms—A case study of the zhanghe river china. *Remote Sens.* **14**(14), 3272, (2022).
21. Fu, B.; He, X.; Yao, H.; Liang, Y.; Deng, T.; He, H.; Fan, D.; Lan, G.; He, W. Comparison of RFE-DL and stacking ensemble learning algorithms for classifying mangrove species on UAV multispectral images. *Int. J. Appl. Earth Obs. Geoinf.* **112**, 102890, (2022).
22. Woebbecke, D. M.; Meyer, G. E.; Von Bargen, K.; Mortensen, D. A. Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE.* **38**(1), 259-269, (1995).
23. Marques, M. G.; da Cunha, J. P. A. R.; Lemes, E. M. Dicamba injury on soybean assessed visually and with spectral vegetation index. *AgriEngineering.* **3**(2), 240-250, (2021).
24. Dai, J.; Xue, J.; Zhao, Q.; Wang, Q.; Cheng, B.; Zhang, G.; Jiang, L. Extraction of cotton seedling information using UAV visible light remote sensing images. *Trans. Chin. Soc. Agric. Eng.* **36**(4), 63-71, (2020).
25. Hamuda, E.; Glavin, M.; Jones, E. A survey of image processing techniques for plant extraction and segmentation in the field. *Comput. Electron. Agric.* **125**, 184-199, (2016).
26. Meyer, G. E.; Neto, J. C. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* **63**(2), 282-293, (2008).
27. Meyer, G. E.; Neto, J. C.; Jones, D. D.; Hindman, T. W. Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Comput. Electron. Agric.* **42**(3), 161-180, (2004).
28. Hague, T.; Tillett, N. D.; Wheeler, H. Automated crop and weed monitoring in widely spaced cereals. *Precis. Agric.* **7**, 21-32, (2006).
29. Liu, Y.; Gong, Y.; Duan, B.; Fang, S.; Peng, Y. Combining multi-temporal NDVI and abundance from UAV remote sensing data for oilseed rape growth monitoring. *Geomatics Inf. Sci. Wuhan Univ.* **45**(2), 265-272, (2020).

30. Kataoka, T.; Kaneko, T.; Okamoto, H.; Hata, S. Crop growth estimation system using machine vision//Proceedings 2003 IEEE/ASME international conference on advanced intelligent mechatronics (AIM 2003). *IEEE*. **2**, b1079-b1083 vol. 2, (2003).
31. Hunt, E. R.; Cavigelli, M.; Daughtry, C. S.; Mcmurtrey, J. E.; Walthall, C. L. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* **6**, 359-378, (2005).
32. Verrelst, J.; Schaepman, M. E.; Koetz, B.; Kneubühler, M. Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. *Remote Sens. Environ.* **112(5)**, 2341-2353, (2008).
33. Gitelson, A. A.; Viña, A.; Arkebauer, T. J.; Rundquist, D. C.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **30(5)**, 1249, (2003).
34. Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M. L.; Bareth, G. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinf.* **39**, 79-87, (2015).
35. Gao, Y.; Lin, Y.; Wen, X.; Jian, W.; Gong, Y. Vegetation information recognition in visible band based on UAV images. *Trans. Chin. Soc. Agric. Eng.* **36(3)**, 178-189, (2020).
36. Zhao, J.; Yang, H. B.; Lan, Y. B.; Lu, L. Q.; Jia, P.; Li, Z. M. Extraction method of summer corn vegetation coverage based on visible light image of unmanned aerial vehicle. *J. Agric. Machinery.* **50(5)**, 232-240, (2019).
37. Zhou, Y.; Zhu, Y. Vegetation extraction from urban aerial images based on visible light. *Green technology.* **(16)**, 247-250, (2016).
38. Louhaichi, M.; Borman, M. M.; Johnson, D. E. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. *Geocarto Int.* **16(1)**, 65-70, (2001).
39. Mao, W.; Wang, Y.; Wang, Y. Real-time detection of between-row weeds using machine vision//2003 ASAE Annual Meeting. *American Society of Agricultural and Biological Engineers.* **1**, (2003).
40. Kawashima, S.; Nakatani, M.; An algorithm for estimating chlorophyll content in leaves using a video camera. *Ann. Bot.* **81(1)**, 49-54, (1998).
41. Guijarro, M.; Pajares, G.; Riomoros, I.; Herrera, P. J.; Burgos-Artizzu, X. P.; Ribeiro, A. Automatic segmentation of relevant textures in agricultural images. *Comput. Electron. Agric.* **75(1)**, 75-83, (2011).
42. Guerrero, J. M.; Pajares, G.; Montalvo, M.; Romeo, J.; Guijarro, M. Support Vector Machines for crop/weeds identification in maize fields. *Expert Syst. Appl.* **39(12)**, 11149-11155, (2012).
43. Sellaro, R.; Crepy, M.; Trupkin, S. A.; Karayekov, E.; Buchovsky, A. S.; Rossi, C.; Casal, J. J. Cryptochrome as a sensor of the blue/green ratio of natural radiation in Arabidopsis. *Plant physiol.* **154(1)**, 401-409, (2010).
44. Woebbecke, D. M.; Meyer, G. E.; Von Bargen, K.; Mortensen, D. A. Color indices for weed identification under various soil, residue, and lighting conditions. *Trans. ASAE.* **38(1)**, 259-269, (1995).
45. Meyer, E. G.; Neto, C. J. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* **63(2)**, 282-293, (2008).
46. Breiman, L. Random forests. *Mach. Learn.* **45**, 5-32, (2001).
47. Yue, J.; Tian, Q.; Dong, X.; Xu, N. Using broadband crop residue angle index to estimate the fractional cover of vegetation, crop residue, and bare soil in cropland systems. *Remote Sens. Environ.* **237**, 111538, (2020).
48. Morais, R. M.; Pedro, J. Machine Learning Models for estimating quality of transmission in DWDM networks. *J. Opt. Commun.* **10(10)**, D84-D99, (2018).
49. Ogunleye, A.; Wang, Q. G. XGBoost model for chronic kidney disease diagnosis. *IEEE ACM T. comput. Bi.* **17(6)**, 2131-2140, (2019).
50. Sampson, J. R. Adaptation in natural and artificial systems (John H. Holland). **18(3)**, 529-530, (1976).
51. Schaffer, J. D.; Multiple objective optimization with vector evaluated genetic algorithms//Proceedings of the first international conference on genetic algorithms and their applications. *Psychology Press.* 93-100, (2014).

52. Hajela, P.; Lin, C. Y. Genetic search strategies in multicriterion optimal design. *Struct. Multidiscipl. Optim.* **4**(2), 99-107, (1992).
53. Handl, J.; Knowles, J. An evolutionary approach to multiobjective clustering. *IEEE Trans. Evol. Comput.* **11**(1), 56-76, (2007).
54. Ghaemi, R.; bin Sulaiman, M. N.; Mustapha, N.; Ibrahim, H. Improving of Initial Clusters Fitness in Genetic Guided-Clustering Ensembles//2010 Seventh International Conference on Information Technology: New Generations. *IEEE*. 227-232, (2010).
55. Liu, Y.; Diao, X.; Cao, J.; Zhou, X.; Shang, Y. A method for entity resolution in high dimensional data using ensemble classifiers. *Math. Probl. Eng.* **2017**, 1-11, (2017).
56. Yao, X.; Fischer, M.; Brown, G. Neural network ensembles and their application to traffic flow prediction in telecommunications networks//IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No. 01CH37222). *IEEE*. **1**, 693-698, (2001).
57. Fernández, A.; Carmona, C. J.; Jose del Jesus, M.; Herrera, F. A Pareto-based ensemble with feature and instance selection for learning from multi-class imbalanced datasets. *Int. J. neural Syst.* **27**(06), 1750028, (2017).
58. Ekbal, A.; Saha, S. A multiobjective simulated annealing approach for classifier ensemble: Named entity recognition in Indian languages as case studies. *Expert Syst. Appl.* **38**(12), 14760-14772, (2011).
59. Onan, A.; Korukoğlu, S.; Bulut, H. A hybrid ensemble pruning approach based on consensus clustering and multi-objective evolutionary algorithm for sentiment classification. *Inform. Process. Manag.* **53**(4), 814-833, (2017).

Acknowledgements

This research was supported in part by the 2022 Doctoral Research Initiation Fund of Hunan University of Chinese Medicine under Grant 0001036, the Hunan Provincial Department of Education Scientific Research under Grant 22B0376.

Author contributions statement

Methodology, H.G., Y.C. and J.L.; software, H.G. and Z.W.; investigation, H.G., Y.C., J.L. and Z.W.; writing—original draft preparation, H.G. and Y.C.; writing—review and editing, H.G., J.L., and Z.W. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.