Thermal Image Diseases Identification Using Hybrid Genetic Algorithm with Relevance Vector Machine Classification

Kesavan Anitha1*, Subramanian Srinivasan2

1 Department of Electronics and Communications Engineering, Saveetha School of Engineering, Saveetha Nagar, Thandalam, Chennai, Tamil Nadu 600077, India
2 Institute of Bio-Medical Engineering, Saveetha Institute of Medical and Technical Sciences. Saveetha Nagar, Thandalam, Chennai, Tamil Nadu 600077, India

Corresponding Author Email: anithasureshbabu241@gmail.com

https://doi.org/10.18280/ts.390530

Received: 11 August 2022
Accepted: 15 October 2022

Keywords:
digital infrared thermal image, hybrid genetic algorithm with relevance vector machine

ABSTRACT

For thermal-based applications, images are obtained by an Infrared camera through Plank’s law. Thermal sensitivity is the smallest temperature difference detected by the camera. Thermal-images were captured through the heat emitted from plant leaves. Initially, the Gaussian noise in the medical Infrared (IR) images is pre-processed by the median filter. Then the features from the preprocessed images are obtained through the principal component analysis algorithm. From the extracted features, the optimal features are selected using the Scale Invariant Differential Evolution-based Feature (SIDEF) algorithm. Through the exploitation of the selected features, the hybrid genetic algorithm with Relevance Vector Machine (HGRVMA) classifier classifies the features into diseases and non-diseases. To validate the performance of the proposed algorithm, it is compared with the existing algorithms in terms of metrics such as sensitivity, accuracy, precision, and recall. The validation results prove that the proposed HGRVM algorithm is optimal than the existing algorithms for all the metrics.

1. INTRODUCTION

Image processing has played a vital role in real-time applications across various fields, such as medicine, computer vision, natural resource management, disaster management, etc., in recent years. The image acquisition and the type of images vary from one application to the next [1]. Image processing is a widespread technique in the field of plant leaves to diagnose a disease or a disruption in the plants. Plant recognition is an area of research work that has existed for a couple of decades. Techniques to recognize the plant automatically from leaf images with the use of computers and pattern recognition flourishes in the 21st century onwards are discussed. Computer vision is a subset of image processing. A computer vision system attempts to simulate vision at the human scale using image processing methods.

The manual identification process requires botanists to have top-to-bottom knowledge of the world's herbaria. This procedure is time-consuming, and in most cases, it can be performed purely by botanists with deep knowledge and who have specialized knowledge in plant taxonomy [2]. Thus, automated tools help to identify and classify a plant, which is a current urgent need in the botanical field. Natural resources, when appropriately managed, serve as the cornerstone for preserving and enhancing the standard of living for everyone on the planet and can significantly contribute to sustainable growth. Environmental protection has triggered an awareness of saving these plants, and botanists are discovering methods to protect them. Out of these methods, using leaves for identifying and recognizing plants has become the most important method. Right now, the vast majority of the current frameworks are utilized for this purpose, which is relevant to certain species and requires human (botanist) intervention to define terms for feature extraction and preprocessing. The practice of efficiently anticipating and responding to calamities is known as disaster management. In order to reduce the damage caused by disasters, resources must be strategically organized and systematic. Medical image processing's key advantage is that it enables thorough, non-invasive investigation of internal anatomy. The process of removing useful information from medical photographs involves frequently applying computational techniques. This emphasizes the fact that the classification system should focus not only on the classification algorithm but also equal attention should also be given to the other phases like pre-processing, feature extraction, and so forth [3].

In agribusiness, the sickness of the executives is the act of limiting illness in harvests to expand the amount or nature of reap yield. The six significant sickness standards the executives are rejection, destruction, security, opposition, treatment, and aversion of bug vectors and weed. Control of plant illness is fundamentally solid on the indiscriminate utilization of substance pesticides, including bactericides, fungicides, and insect poisons that are hurtful for plant microorganisms or plant infection vectors. Among these control techniques, obstruction is one of the most monetary and eco-accommodating strategies to control plant sickness. Illness is the board that engages people, working with other medical care suppliers to deal with their infections and forestall entanglements. Sickness of the executives has arisen as a promising technique for further developing consideration for those people with constant circumstances. Boiling water
treatment is generally utilized for the control of seed-borne microbes, particularly microscopic organisms.

A plant temperature that is unusual is a sign of disruption. Using a digital infrared thermal image (DITI) to monitor temperature variation is a valuable and non-invasive method. Infrared imaging systems provide lofty-resolution images of plant leaves temperature and can be accustomed to calculate responsive changes in leaf structure temperature in regard to the abnormalities and their changes to the prevention of plant diseases [4]. FLIR Thermal cameras with a sensitivity of 0.01°C with a temperature range from -200°C to +1200°C are used to capture thermal images. Advanced image processing techniques that make use of thermal images have a wide variety of real-time applications such as UVA surveillance, pedestrian detection, non-destructive testing, Volcano logy, fault detection, military applications, and in most of the security-based systems [5]. Infrared-based image regions are classified into four types, as shown in Figure 1.

Because of the low attenuation losses in the SiO$_2$ glass (silica) medium, this wavelength range widely utilized in fiber optic telecommunications goes up to the wavelength of the first water absorption band. This region of the spectrum is sensitive to image intensifiers. The use of near-infrared spectroscopy is another typical example. At 1,450 nm, there is a noticeable increase in water absorption. The predominant spectral region for long-distance communications is in the range of 1,530 to 1,560 nm. In guided missile technology, the homing heads of passive IR "heat seeking" missiles are designed to function in the 3-5 m region of this band. The "warm imaging" district, in which sensors can get a detached picture of items just somewhat higher in temperature than room temperature for instance, the human body-in view of warm outflows requiring no enlightenment like the sun, moon, or infrared illuminator. This area is likewise called the "warm infrared." Far-infrared laser or terahertz laser is a laser with a yield in the middle of between 30-1000 μm in the far infrared and terahertz recurrence band of the electromagnetic range.

In this research, the IR thermal image concept and hybrid classification method were applied to plant disease detection. The investigation and identification procedure have to address this state of affairs carefully. Accurate and efficient characteristic extraction strategies that greatly distinguish these similar leaves are required for a successful design of the automated framework. Further, the availability of a huge number of leaf features and selecting a subset that amazingly enhances the technique of identification is hard [6]. Figure 2 explains the general architecture diagram for thermal leaves diseases detection.

There is no generalized methodology that can detect all kinds of abnormalities. Segmentation: The Selection of ROI is the challenging factor. Feature Extraction: The existing models are highly complex in nature and do not handle the data efficiently. It has a low accuracy rate and less ability to extract the features. Feature Selection: The present techniques retain more memory, and the computation of numerical value is complex [7]. Image Classification: the transparency of the result is low, and computation complexity occurs in the surviving technique. From the analysis of the existing detection algorithms, it is clear that they do not provide a satisfactory accuracy rate.

Further, the existing thermal image processing algorithms have higher tracking errors, non-optimal performance, and higher complexity [8]. The proposed hybrid HGRVMA technique occasionally builds the arrangements for efficiency; accuracy and error rate are nearly regular strategies. The HGRVMA technique accomplished higher proficiency in 90.15% than in other ordinary strategies, KNN is 86.45%, PCA is 83.18%, and SVM is 78.64% accomplished. The proposed HGRVMA calculation accomplished 3.70% higher than the most noteworthy ordinary k-nearest neighbor's calculation.

The remainder of this research is organized as chapter 1 presents an introduction to the plant kingdom, automatic plant identification task along with the research objectives. Section 2, The literature review is a critical examination of previous research that is relevant to the topic and is completed. Section 3 discusses to Hybrid genetic algorithm with the Relevance Vector Machine; section 4 presents the Hybrid genetic algorithm with the Relevance Vector Machine system and conventional experimental results comparison. Finally, section 5 provides the concluding remarks and future scope of the work.

2. LITERATURE REVIEW

This chapter discusses the uses of thermal images in a variety of fields, with a focus on the thyroid and diabetic foot. Finally, a three-phase survey is conducted after a briefly introduction to thermography and its applications. The various techniques involved in thermal image processing are presented
first, followed by a survey of the methodologies used by various researchers in abnormality detection and classification. Finally, the importance of thermal images in the agricultural field is surveyed, resulting in the identification of research gaps and the formulation of research objectives.

Contour-Based Registration is used to register images from different sources and combines different features [9]. Extraction of gradient information from the image, where the value of each pixel represents the boundary of a foreground objects [10]. Affine registration sufficiently matches the images of a scene taken from the same view but from different position [11]. Feature based Registration handles the complexity between image distortions and relies on a small number of features [12] and abstracts significant data from the original data input and filters out the redundant data [13]. The watershed transform is a widely used district-based picture segmentation algorithm that takes its strategy from topography [14]. Warm balance in the feet is an ordinary tracking down in sound subjects. It is fascinating that a few investigations have been performed with minimal-expense radiation thermometers. The patients were told to self-screen by estimating temperatures on the incredible toe; first, third, and fifth metatarsal heads; mid foot; and impact point consistently. In light of the application, a solitary or a blend of division systems can be connected to take care of the issues successfully [15].

The Correlation-based Feature Selection (CFS) method is used for CT images to determine the best feature subset and computes the correlation between two random variables. The method uses rigid translation and rotation of an image [16]. It automatically selects optimum predictors at various levels of model complexity specified by the system in [17] using CFS for breast and heart datasets. PSO selects more efficient subsets and is used to generate decision rules for degree prediction, which increases the prediction rate [18]. A framework has been designed with Hybrid GA that allowed using high-level textural features. The Fitness value is determined by the match between the textures [19]. At this point, they used pelvic computed demographics images to obtain the result [20]. Global neighborhood Structure (GNS) makes use of considering the intensity-based similarities among the pixels in an image. The work is concentrated on the texture feature of an image [21]. A system was developed for brain images, and the system efficiently classified the human brain as Malignant or Benign [22]. A hybrid Genetic Algorithm emulates chromosomes, reproduction, and selection used to create a population of individuals [23]. Particle Swarm Optimization (PSO) is applied to a medical dataset in which each particle uses its own memory, and optimal solution and solves discrete problems. The classifier reduces the noise and determines the property of the fuzzy sets and fuzzy rules [24]. ANF classifier work is carried out on the cardiac dataset, and the result is obtained [25]. A multilayer Feed Forward Network (FNN) is utilized in which each node performs a particular function. Here in this work, the author has computed in designing the mean square error curve for the classifier [26]. The decision Tree based Classification method handles the problem of over fitting the data and generalizes the tree by removing the noise and outliers. It requires more time to build the model for prediction [27], toe; first, third, and fifth metatarsal heads; mid-foot; and impact point consistently. They were told to call the review nurture on the off chance that a temperature contrast of 4 °F was found.

Skin temperature, the data that can be explored by this procedure, relies upon blood flow. The strange internal heat level is a characteristic sign of sickness. Infrared thermograph (IRT) is a quick, detached, non-contact, and harmless option in contrast to customary clinical thermometers for observing internal heat levels. Warm pictures likewise experience the ill effects of a nearly low sign-to-commotion proportion (SNR). Consequently, picture handling is of prime significance in the field of IRT. Different channels (in both time and recurrence spaces) and calculations are utilized for limiting commotion, diminishing obscuring, and edge safeguarding in warm pictures. SVM constructs a hyper plane in a high or infinite dimensional space and works well for binary classification, and has a higher accuracy rate in classifying the data. Adaptive Neuro-Fuzzy Classifier has been used as a Neuro based classifier model which discriminates the normal and abnormal patterns from the brain region.

A fuzzy rule has been generated based on the features, and an inference system has been designed [28]. Distinguishing solid biomarkers for anticipating thyroid knob finding needs comprehension of all parts of disease cell demise and survival. The thyroid knob tissue tests were recuperated from80 patients who had experienced surgical treatment. The fractional Gal-3 quality variety as an analytic marker for thyroid knobs speaks to a promising road for future study, and its clinical application could fundamentally decrease the number of symptomatic thyroid operations performed for instances of determinant fine thermal images [29] and an investigation is suggested in reference [30] that improve the accuracy of breast mass conclusions (malignant or benign). The findings of four clinical cases (female patients with malignant tumors and benign tumors) show that, regardless of breast density, the estimated values of the power of heat sources in malignant instances are significantly more apparent than those in benign cases. The correlation coefficients (R2) of the nonlinear curve fittings are all above 0.98. All the nonlinear curve fits’ correlation coefficients (R2) are more than 0.98. A novel tool for farmers to automatically detect plant leaf diseases is discussed in the study of [31]. The diseased leaf spots are first identified using fuzzy c-means clustering. Gray-level co-occurrence matrix is used to extract the features, and progressive neural architecture search is used to categories them.

The calculations that distinguish and sort leaf sicknesses are not extremely exact and productive. Precision can be improved to forestall affliction and disarray. Machine learning techniques for sickness evaluation consume most of the day and require a great deal of preparing information. Image handling's true capacity for exact disease seriousness appraisal has not been totally researched.

Numerous pattern recognition and object identification applications are generated, according to the literature, but the feature retrieved is insufficient to distinguish the object from the image. Computer-assisted plant leaf classification is a difficult undertaking that involves the following concerns. (1) A method for identifying plants in general (2) having to deal with a large variety of different plant species (3) the segmentation of appealing characteristics including shape, texture, and colour (4) Choosing a collection of characteristics that is a good predictor of which group a sample belongs to.

3. SYSTEM DESIGN

This chapter states the implementation of computerized
techniques on plant leaves diseases identification and Classification from Thermal image Implementation done using GLCM feature extraction [32] and Hybrid genetic algorithm with Relevance Vector Machine methods for Classification. Finally, results are compared and analyzed against performance metrics.

The temperature distribution or the measurements of the thermal pattern recorded in an infrared thermal image are from the external surface layer of the plant leaves and tends. The emissivity of plant leaves is independent; it has also been reported that the plane curve plays a vital role in the value of plane emissivity. DITI is successfully used in the analysis of other applications for fever screening, satellite images, heart functioning, breast cancer, diabetes, thyroid disorder, thyroid eye, plant diseases, knee injury, and peripheral vascular disorders. Thermal imaging has arises to provide target estimation of temperature changes that are clinically critical in certain therapeutic application. With basic utilization of the innovation and appropriate comprehension of thermal physiology, Thermal imaging us a tenable and acceptable diagnosis tool in medicine and remedies. The camera's thermal sensitivity is the lowest temperature difference it can detect. For example, the temperature sensitivity of an infrared camera with an uncooled micro bolometer detector is 0.010C at 300C; at 300°C, an infrared camera with a semiconductor type of detector has a temperature sensitivity of 25mK. Infrared emissions from the human plant architectural membrane at 27°C defamation, with a wavelength range of 2–20 lm and a peak of around 10 lm. Body infrared rays are a wavelength band with a significantly shorter wavelength (8–12 lm) used for medical purposes. For wavelengths between 2 and 14 lm, the emissivity of the human skin membrane is 0.98 0.01. Infrared energy from an object is converted into images by a thermal camera (thermograms). A source of radiation is an object's exterior temperature and emissivity. The ambient absorption affects the incoming emission from the object. An approach is developed for the detection of plant abnormality using a Hybrid genetic algorithm with Relevance Vector Machine methods.”

![Figure 3. Electromagnetic spectrum](image)

The electromagnetic Spectrum from Figure 3 shows section 1 belongs to X-ray, 2 is ultra violet ray, 3 is visible light, 4 is an infrared region, 5 is microwave, and 6 is a radiofrequency band. Figure 4 explains the Hybrid genetic algorithm with a Relevance Vector Machine architecture diagram for thermal plant leaves image classification.

Digital plant leaves consist of a vast amount of knowledge that predicts the original tissue, ducts, lumps, and breast edges, in view of developing a robust diagnosis system for classifying it as diseased or not diseased [33]. In our approach, we examined 22 features applied to the region of interest of window size 75 pixels with 75 pixels shift without any overlapping. Texture investigation is an essential area of study in thermal imaging and processing the images with a set of procedures. Texture analysis constitutes of three most important kinds of problems, which consist of texture classification, segmentation, and synthesis. In general, synthesis is used in image compression.

![Figure 4. Thermal plant disease detection using hybrid genetic algorithm with relevance vector machine classification](image)

Here, the Figure 4 model explains with work procedure: the genetic algorithm gathers the various inputs, and further feature values are applied to the support vector machine to select the appropriate rule. Here we have framed about 64 rules and classified them as diseased, Non-diseased, and May be diseased depending on the values of Contrast, Homogeneity, Average of Features obtained from the Co-occurrence Matrix, and Expert opinion obtained from a Radiologist. Instead of numeric values for each feature, ranges have been set to indicate the severity of the features as L: Low, M: Medium, and H stands for High.

A successful leaf identification system requires a set of leaf features that best portray the leaf image, a picture dataset separated into various classes, and a picture classifier to be made. After the information result explored the dataset, the K-closest neighbor (KNN) is by all accounts a decent choice. To utilize the KNN calculation, there is a significant boundary to utilize, which is K which can give the most extreme segregation between different levels. The collection of features of an image is represented as a pattern. The features are the characteristic of the training images for a given problem. For an object recognition system, the features might be local or global properties extracted from the image. Classifiers are classified into two bases: they are discriminatory and generative models. The discriminatory model learns directly from the training data. The discriminative model produces a prediction function with given constraints and parameters using the training samples. In the generative model, learning through a statistical model is created, which can predict the unknown to be known. The work cycle of GA is described as making the principal
population of individuals and evaluating every individual's wellness inside that segment. Proceed with this age until it closes. Pick the individuals who will profit from hereditary activities the most. Utilize the hereditary administrator’s hybrid and change to create the ideal gathering of descendants. Survey the individual wellness of new individuals from the posterity bunch. Pick the top individuals for the accompanying age.

Real-time classification problems are challenging to solve, and many applications have to pact with Non-deterministic Polynomial (NP)-hard problems. To sort such problems, classification tools have to be used, though the optimal solution obtained, the proposed system optimizes the weight of the GA structure, using the Relevance Vector Machine (RVM) proposed efficiently.

4. RESULT AND DISCUSSION

The methodology presented in this work was tested on the complete Thermal images. It is available for research purposes in DITI. The images of the database originated from a FLIR Thermal camera and were processed as a screening technique. The algorithms were implemented based on 324 trained images consisting of 176 normal and 148 abnormal plant leaves. The research shows various features using GLCM and sensitivity, specificity, and accuracy values by various classification algorithms from the GLCM feature extraction algorithm. The plant database consists of 324 images. The images are arranged in pairs. The plant (even file numbers) has the right plant leaves (odd file numbers) of a single leaf, and the database has 148 abnormalities.

![Figure 5. Leaf thermal images dataset with diseases classification](image)

The types of abnormalities present in plant leaves are microclassification, masses, architectural distortion, and bilateral asymmetry. If the classification is present, center location and radius are applied to the cluster rather than the individual classification. For this study, a total of 322 plant leaves well taken have been digitized to a resolution of 200-micron pixel edge and padded. Figure 5 explain the various diseases and corresponding thermal image pattern. Table 1 explains the performance Comparison for the following false positive, True positive, and F-score.

<table>
<thead>
<tr>
<th>Methods</th>
<th>False positive</th>
<th>True positive</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support Vector Machine</td>
<td>26.18</td>
<td>73.61</td>
<td>49.44</td>
</tr>
<tr>
<td>Personal Component Analysis</td>
<td>38.59</td>
<td>76.81</td>
<td>56.72</td>
</tr>
<tr>
<td>k-nearest neighbors algorithm</td>
<td>49.79</td>
<td>82.49</td>
<td>66.69</td>
</tr>
<tr>
<td>Proposed HGRVMA system</td>
<td>76.28</td>
<td>86.14</td>
<td>79.22</td>
</tr>
</tbody>
</table>

![Figure 6. Performance comparison for thermal image plant diseases dataset](image)

<table>
<thead>
<tr>
<th>Methods</th>
<th>Efficiency</th>
<th>Precision</th>
<th>Error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support Vector Machine</td>
<td>78.64</td>
<td>81</td>
<td>88</td>
</tr>
<tr>
<td>Personal Component Analysis</td>
<td>83.18</td>
<td>86</td>
<td>91</td>
</tr>
<tr>
<td>k-nearest neighbors algorithm</td>
<td>86.45</td>
<td>87.65</td>
<td>94</td>
</tr>
<tr>
<td>Proposed HGRVMA system</td>
<td>90.15</td>
<td>91.11</td>
<td>76</td>
</tr>
</tbody>
</table>

![Figure 7. Efficiency comparisons for thermal image plant diseases dataset](image)
Figure 6 shows the graphical representation for false positive, True positive, and F-score. In terms of thermal plant leaves image GLCM attributes. The proposed hybrid HGRVMA method periodically increases the sequences for false positive, True positive, and F-score are comparatively conventional methods. Table 2 explains the performance comparison in terms of efficiency, precision, and error rate. Figure 7 shows the graphical representation for efficiency, Precision, and Error rate.

The proposed hybrid HGRVMA method periodically increases the sequences for False Efficiency, Precision, and Error rate are comparatively conventional methods. HGRVMA method achieved higher efficiency at 90.15% in other conventional methods, KNN is 86.45%, PCA is 83.18%, and SVM is 78.64%. The proposed HGRVMA algorithm achieved 3.70% higher than the highest conventional k-nearest neighbor’s algorithm.

Table 3 explains the time performance comparison for the thermal image in plant leave diseases. The duration of prediction and efficiency is a very important parameter for the prediction of leaves disease.

**Table 3.** Time performance comparison for thermal image plant diseases dataset

<table>
<thead>
<tr>
<th>Methods</th>
<th>Time (Ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support Vector Machine</td>
<td>31.25</td>
</tr>
<tr>
<td>Personal Component Analysis</td>
<td>23.45</td>
</tr>
<tr>
<td>k-nearest neighbors algorithm</td>
<td>18.56</td>
</tr>
<tr>
<td>Proposed HGRVMA system</td>
<td>11.12</td>
</tr>
</tbody>
</table>

Figure 8 shows HGRVMA method achieved a lower time duration for prediction, only 11.12 ms. In other conventional methods KNN is 18.56 ms, PCA is 23.45 msc, and SVM is 31.25 ms, achieved. Comparatively conventional methods, KNN 7.44 ms, PCA 12.33 ms, and SVM 20.13 ms higher time to achieve the corresponding efficiency attained. The proposed system attained the all the quality matrix to better comparatively conventional Support Vector Machine, Personal Component Analysis, and k-nearest neighbors algorithm.

The recognition of five different paddy leaf diseases, Blasts, Bacteria leaf blight, Hispa, Leaf spot, and Leaf Rusts diseases, and their classification performance quality matrices results are briefly discussed in this Chapter. It could be summarized that, for detecting the five different diseases, HGRVMA provides better results. It was proven that the paddy crops cultivated in the self-constructed greenhouse environment were more effective for employing image processing for proposed HGRVMA disease detection than the crops cultivated from the agricultural environment.

### 5. CONCLUSIONS

In precision agriculture, one of the evolving research areas is automated system development for the identification and classification of several rice crop diseases. It may result in the improvement of the quality and quantity of agricultural products. Manual disease identification provides a lack of accuracy for the farmers. Hence, it requires the method of image processing for the accurate, timely disease detection of plants in several cases, as it minimizes the application of human vision. In large farms, the computerized technique of image processing is adapted for the production of rice crops, which uses color information for the detection of diseases on rice leaves. From the reported information from the thesis, an expert system is promoted to rice field farmers for the early recognition of rice crop diseases. The main contributions of this research work are the early detection of Blast, Bacteria leaf blight, Hispa, Leaf spot, and Leaf Rusts diseases the enhancement of overall plant production using thermal images with a Hybrid genetic algorithm with a Relevance Vector Machine. The proposed system is compared to a conventional Support Vector Machine, Personal Component Analysis, and k-nearest neighbor’s algorithm. It follows that use of thermal imaging ideally requires the combined application of one or more other imaging techniques to achieve best results and also a direct execution study of the suggested strategy that encourages quick defense against the previous techniques. The proposed system achieved better efficiency of 90.15% and precision of 91.11% with a minimum time requirement of 11.12 Milliseconds. Other important results the thermal image dataset produce better results than the comparatively normal dataset; thermal dataset clearly differentiates effected and non-affected parts of plant leaves and their efficiency too. In the future need to collect large thermal training datasets, and testing datasets will increase the performance evaluation results.

### REFERENCES


