Jurnal Teknik Informatika C.I.T Medicom, 16 (2) (2024) 58-69

Published by: Institute of Computer Science (IOCS)

Jurnal Teknik Informatika C.I.T Medicom

Journal homepage: www.medikom.iocspublisher.org

Classification of eucalyptus leaves: Combining color histogram feature extraction and decission tree algorithm

Sarifah Agustiani[1], Rahmat Hidayat[2], Yoseph Tajul Arifin[3], Haryani[4], Siti Marlina[5]

[1),2),3),4)] Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Jakarta, Indonesia

[5)] Sistem Informasi, Fakultas Teknologi Informasi, Universitas Nusa Mandiri, Jakarta, Indonesia

Article Info

Article history

Received : Mar 20, 2024
Revised : Apr 13, 2024
Accepted : May 18, 2024

Keywords:

Classification;
Color Histogram;
Decision Tree;
Eucalyptus.

Abstract

This research proposes an automatic approach to identify eucalyptus species based on leaf images using color histogram feature extraction and the Decision Tree algorithm. Eucalyptus is known as one of the most productive plants in the world with various uses in the timber, biofuel and pharmaceutical industries. However, its wide environmental adaptability and rapid growth pose challenges in identification and management. The proposed approach focuses on the use of Artificial Intelligence (AI) technology and image analysis to solve the identification problem. The color histogram feature extraction method is used to extract visual information about the color distribution of eucalyptus leaves. The Decision Tree algorithm is used to build a classification model based on the extracted features. Model evaluation is carried out using accuracy, precision, recall and F1-score metrics. The results showed that this approach was effective in identifying eucalyptus species, with a high level of accuracy. In addition, the development of this method offers opportunities for further applications in various fields, including forest mapping, mobile applications, and the timber industry. By combining advances in AI and image analysis, this research has the potential to become an important cornerstone of nature conservation and environmental sustainability efforts, and help strengthen natural resource management globally.

Corresponding Author:

Sarifah Agustiani,
Program Studi Teknologi Komputer,
Universitas Bina Sarana Informatika,
Jl. Kramat Raya No.98, Kec. Senen, Daerah Khusus Ibukota Jakarta 10450, Indonesia,
Email: sarifah.sgu@bsi.ac.id

1. Introduction

The practice of clonal forestry on eucalyptus has made this plant one of the most productive plants in the world [1][2]. Eucalyptus not only provides the main raw material for the paper, wood and biofuel industries, but also contributes to the medical industry through the extraction of eucalyptus oil or eucalyptol which has antiviral and anti-inflammatory properties [3][4]. Eucalyptus is also known for its broad environmental adaptability, rapid growth, and diverse uses, ranging from pulp and paper production to bioenergy and essential oil extraction [5]. However, this richness of species also poses challenges in terms of identification [6].

The urgency of the problem in Indonesia is increasing along with industrial growth and the need for natural resources [7]. Although eucalyptus has significant benefits in various industries, its rapid growth and wide adaptation can result in serious environmental problems if not managed properly [8]. Each species has characteristics that lead to different exploitation strategies and industrial uses. Therefore, current knowledge of the abundance and spatial distribution of various species is important for forest planning [9]. One of the main problems is the risk of invasion by alien species that could disrupt native ecosystems and threaten local biodiversity [10].

Eucalyptus plants that grow abundantly in many parts of Indonesia can compete with native plant species, changing the composition of local flora and reducing biodiversity [11]. Apart from that, improper management of eucalyptus forests can also cause land degradation, soil erosion and reduce groundwater quality [12]. In this context, it is important to have effective methods for identifying and managing eucalyptus plants appropriately in order to minimize their negative impact on the environment.

Research that combines Artificial Intelligence (AI) technology with leaf image analysis can provide innovative solutions in overcoming identification problems [13] including identification of eucalyptus species in Indonesia. Artificial Intelligence (AI) is a field in computer science that focuses on developing computer systems that can perform tasks like humans [14]. Artificial Intelligence is able to imitate human behavior where all its actions are considered intelligent or smart [15]. One important aspect of this development is progress in the field of Computer Vision [16]. Computer Vision has become an integral part of AI, enabling computers to understand and analyze visual data such as images and videos with a level of intelligence that is increasingly approaching human capabilities [17]. Various technologies in Computer Vision, such as object recognition, face detection [18][19], image segmentation [20]–[22], even in the field of medicine for medical image-based disease diagnosis [23]–[29]. Using automated methods to classify eucalyptus species based on visual characteristics of leaves, such as shape, texture, vein pattern, and color, can help foresters and land managers identify species more efficiently and accurately.

In the context of developing methods for classifying eucalyptus types based on leaf images used in this research are the color histogram feature extraction technique and the decision tree algorithm. The color histogram feature is one of the approaches commonly used in image analysis to extract information about the color distribution in an image [30]. By measuring the frequency of occurrence of each color value in the image, this feature can provide a fairly strong representation of the visual characteristics of leaves [31] including eucalyptus leaves.

The decision tree algorithm, on the other hand, is a machine learning technique that is capable of learning patterns from training data and making decisions based on the learned rules [32]. The same method is also applied in several studies [9], [33]–[37]. In the context of eucalyptus species classification, the decision tree algorithm can be used to build a model that can differentiate between eucalyptus species based on features extracted from leaf images. The main advantage of the decision tree algorithm is its ability to produce models that are easy to interpret [38]–[41], thereby enabling forest researchers and practitioners to understand the reasons behind each classification decision.

By combining the color histogram feature extraction technique to extract visual information from eucalyptus leaf images and a decision tree algorithm to build a classification model, this research can provide a powerful and effective approach in automatically identifying eucalyptus species. The combination of these two techniques can produce a system that is able to overcome the challenges of identifying eucalyptus species accurately and efficiently, and can be used as a valuable tool in forest management and biodiversity conservation.

Thus, this research will not only contribute to protecting Indonesia's biodiversity and managing natural resources sustainably, but will also facilitate the development of a more responsible and environmentally friendly eucalyptus industry. Through the application of AI technology in the forestry sector, it is hoped that we can utilize the potential of eucalyptus optimally while still paying attention to environmental protection and sustainability of local ecosystems.

2. Research Methods

This research methodology consists of several stages which can be seen in Figure 1. The first stage involves collecting images of eucalyptus leaves obtained from the Kaggle repository. Next, color histogram features were extracted from each image to capture visual information about leaf color distribution. The next stage is the application of the decision tree algorithm to build a classification model based on the extracted features. After the model is built, testing and evaluation is carried out using separate test data to assess accuracy, precision, recall and F1-score. Thus, this research combines image feature extraction techniques and machine learning in an automatic approach to classify eucalyptus types based on leaf images. Figure 1 shows these stages visually to provide a clearer picture of the methodology of this research.

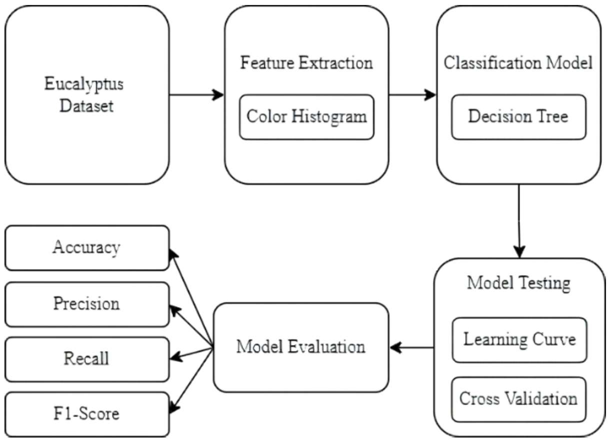

Figure 1. Research Method

a. Dataset

This research uses a eucalyptus dataset taken from the Kaggle Repository [42]. The dataset is divided into three classes, namely the Cirtriodora class with 281 images, Grandis with 311 images, and Robusta with 329 images. Thus, the number of images in the dataset used for this research is 921 images.

b. Feature Extraction

Feature extraction is the process of identifying, selecting, or extracting specific information from images or visual data for use in further analysis or in model development [43]. In the context of color histograms, this extraction feature focuses on the distribution of colors in an image.

Color histogram is a feature extraction method commonly used in image processing [30]. Color histograms are used to measure how often different levels of color intensity appear in an image. Basically, it involves grouping color values into intervals or bins and counting the number of pixels that fall into each bin. This histogram provides a numerical representation of the color distribution in an image [44].

c. Decision Tree Algoritm

The Decision Tree algorithm is a special method for the classification process [45], consists of nodes that form a branch structure, representing the decisions produced by the algorithm, and each leaf node represents an outcome [46]. There are two types of nodes, namely decision nodes which have various branches and are used to make decisions, and leaf nodes which are the result of decision nodes without having further branches.

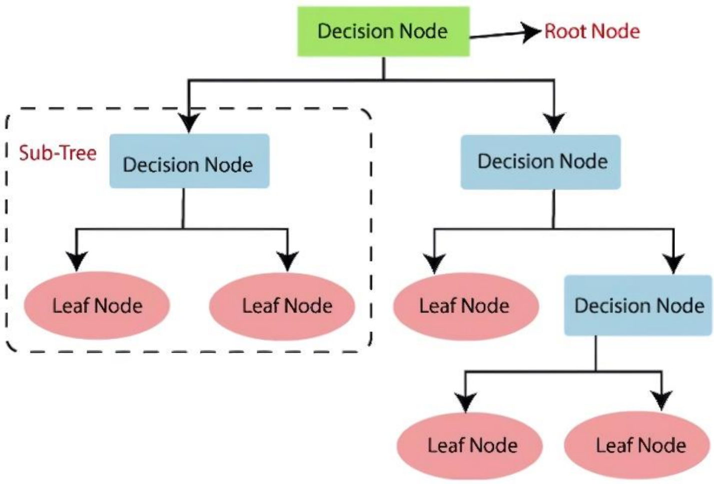

Figure 2. Decision Tree
Sumber: [47]

Terminology for Decision Trees:

1) Root Node: The initial part of a Decision Tree where the entire data set is first divided, forming various homogeneous groups.
2) Leaf Node: The final node where no additional splits are possible in the Decision Tree.
3) Splitting: The process of dividing the main node further, forming sub-nodes based on defined boundaries.
4) Sub Tree: Splits the resulting hierarchy into smaller sub-trees or branches.
5) Pruning: The step of removing excessive branches in the Decision Tree to achieve optimal results. There are two types of pruning: cost complexity and error reduction pruning.
6) Child and Parent Nodes: The base node, also called the parent node, and the other nodes are called child nodes.

Attribute Selection Steps:

The attribute selection process (ASM) involves identifying optimal attributes on source nodes and sub-nodes. There are two main approaches in ASM:

1) Gain Information
Information Gain, as the name suggests, measures the amount of information provided by a class-related feature. Nodes are separated and a tree is built based on the Gain Information values. The Decision Tree algorithm aims to maximize the Information Gain function and prioritize separating nodes or attributes with the highest amount of information first. The Gain Information calculation can be represented mathematically as follows:

$$Information\ Gain = Entropy\ (s) - [(Weighted\ Avg.) * (Entropy(Every\ Feature))] \quad (1)$$

Entropy measures the level of uncertainty in a data set and is represented as follows:

$$Entropy\ (s) = P(yes)Log_2P(yes) - P(no)Log_2P(no) \quad (2)$$

2) Gini Index
The Gini Index measures the level of pollution or purity applied in creating a decision tree. The Decision Tree algorithm prefers attributes with a smaller Gini Index, as this indicates higher purity when making decisions. The calculation of the Gini Index can be stated as follows:

$$1 - \sum_j P_j^2 \quad (3)$$

d. **Testing Model**

1) Learning Curve
A learning curve is a graph that shows how the performance of a model or machine learning algorithm improves as the amount of training data used increases [48]. Learning curves can provide insight into the extent to which a model can gain knowledge or skills from the given data. In a learning curve, the horizontal axis represents the number of training data or iterations, while the vertical axis shows model performance evaluation metrics, such as accuracy or prediction error [49]. Thus it can identify patterns such as underfitting or overfitting.
2) Cross Validation
Cross-validation is a technique in evaluating the performance of a machine learning model that involves dividing a dataset into training and testing subsets repeatedly. The main goal is to mitigate the risk of overfitting and produce more consistent and reliable estimates of model performance [50].

e. **Evaluation Model**

Evaluation of the results of the eucalyptus species classification method using the color histogram feature extraction technique and the decision tree algorithm was carried out by looking at how accurate the model was in identifying eucalyptus species from leaf images. This evaluation involves important metrics such as accuracy, precision, recall, and F1-score. The accuracy value will show how well the model is at predicting overall. Meanwhile, precision and recall will give an idea of how good the model is at distinguishing species correctly. By looking at the F1-score, it can be seen whether the model performance is balanced between precision and recall. This evaluation helps assess the reliability of this method for forest management and biodiversity conservation.

3. Results and Discussion

This study aims to evaluate the effectiveness of automatic classification of Eucalyptus leaves by combining color histogram feature extraction and the Decision Tree algorithm. The application of Color Histogram feature extraction involves the process of analyzing the distribution of color intensity in an image. This extraction process stage begins by dividing the colors in each image into 3 main color channels, namely red, green and blue or what is usually called the RGB color model. Next, a histogram will be formed for each channel by recording the frequency of appearance of each color intensity value. Then a normalization process is carried out by dividing each frequency value by the total number of pixels in each image. This is done to compare images of different sizes, because each channel has a different total number of pixels. After normalization, the histograms from each channel will be combined into one vector in one dimension, this will create a cumulative representation of the color distribution throughout the image. Thus producing a histogram vector that can be used as a feature, the results of this vector provide a strong description of the color characteristics of the image in visual form. Thus, this feature will make it easier for the model to recognize and differentiate images. Figure 1 is one of the histogram results for each class.

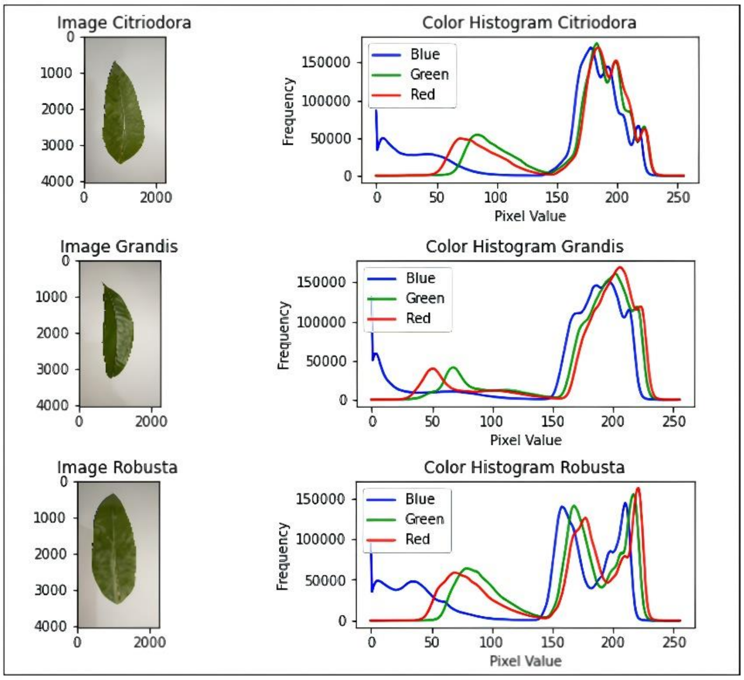

Figure 3. Histogram Results

Figure 3 is the result of extracting color histogram features for each class in this study. The color histogram feature extraction process is an important stage in image analysis, which aims to identify and record the distribution of color intensity in each image. In this context, the image shows a visual representation of the color characteristics taken from images of eucalyptus leaves in each class. Through the use of this feature extraction technique, important information about color patterns and pixel distribution is extracted for each image, allowing the formation of a numerical representation that can be used in subsequent classification processes. Thus, Figure 3 provides a clear picture of how color data is collected and represented in the context of eucalyptus leaf species identification using an image analysis approach.

After getting the feature results from all images through Color Histogram feature extraction, the next step is to carry out the classification process using the Decision Tree algorithm. The Decision Tree algorithm is used to build a decision tree based on the features extracted from each image. This decision tree becomes a classification model that can distinguish and identify various types of Eucalyptus leaves. This classification process utilizes the information contained in the color histogram feature vector as a basis for decision making, providing an automatic and efficient ability to classify leaves into predetermined classes, such as Eucalyptus Citriodora, Eucalyptus grandis, and Eucalyptus robusta.

The application of the Decision Tree algorithm plays an important role in the Eucalyptus leaf classification research in this study. This algorithm works by building a decision tree that groups Eucalyptus leaf images based on features that have been previously extracted using color histograms. The feature selection and thresholding process is carried out to minimize impurities within each node, allowing the decision tree to identify unique patterns that will differentiate between three types of Eucalyptus leaves namely Citriodora, Grandis, and Robusta. By following the rules at each node, this decision tree can classify new leaves into the appropriate class. The following are the results of the decision tree produced from the Decision Tree model in classifying Eucalyptus leaves.

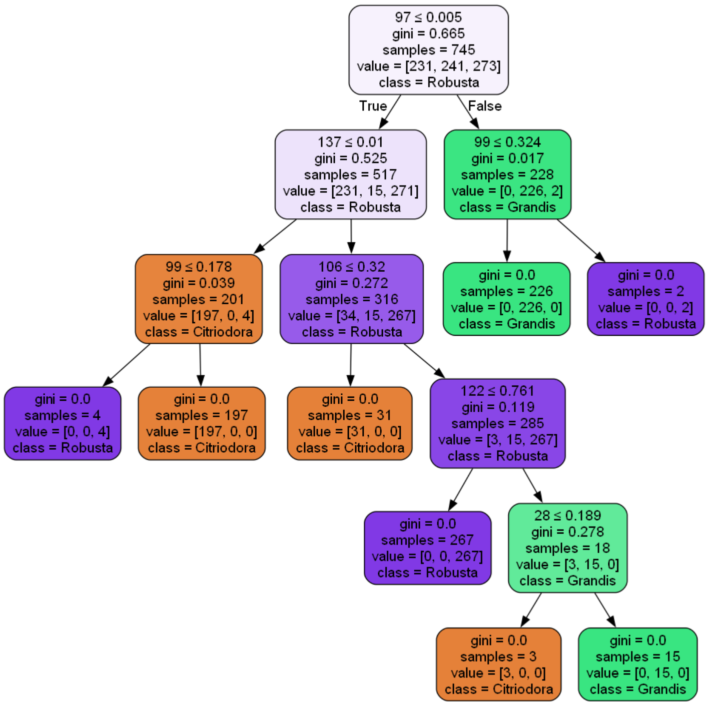

Figure 4. Decision Tree Visualization

Figure 4 is the overall decision tree formed by the model, consisting of nodes that represent decisions based on data features. Each branch in this decision tree represents a decision or split based on certain feature values. The path that data follows from root to leaf reflects a series of decisions that express the predicted outcome of each class. Apart from decision trees, this model also produces rule models which can be seen in Figure 5.

```
Decision Tree Rules:
|--- 97 <= 0.01
|   |--- 137 <= 0.01
|   |   |--- 98 <= 0.60
|   |   |   |--- class: 0
|   |   |--- 98 >  0.60
|   |   |   |--- class: 2
|   |--- 137 >  0.01
|   |   |--- 106 <= 0.32
|   |   |   |--- class: 0
|   |   |--- 106 >  0.32
|   |   |   |--- 122 <= 0.76
|   |   |   |   |--- class: 2
|   |   |   |--- 122 >  0.76
|   |   |   |   |--- 81 <= 0.02
|   |   |   |   |   |--- class: 0
|   |   |   |   |--- 81 >  0.02
|   |   |   |   |   |--- class: 1
|--- 97 >  0.01
|   |--- 44 <= 0.05
|   |   |--- class: 1
|   |--- 44 >  0.05
|   |   |--- class: 2
```

`Figure 5. Rule Model

Figure 5 is a rule model resulting from the proposed model for eucalyptus leaf classification. This rule model is the same as a decision tree, only this rule model refers to individual rules or conditions applied by the model to the decision tree. Each branch and each leaf can be interpreted as a rule that determines how decisions are made based on the values of certain features.

The rule model in Figure 4 shows that the first rule starts by examining a feature in the data (in this example, the 97th feature) with an initial decision based on the feature value (<= 0.01 or > 0.01), then splitting based on the feature, if the feature value initial <= 0.01, the model will examine the 137th feature and if the value of the 137th feature <= 0.01, the model will involve the 98th feature. And so on until it forms a series of branches. So the data will arrive at one leaf which shows the class prediction.

In this way, decision trees break down the complexity of a classification problem into a series of decisions that are simpler and easier to interpret. The features tested by the model form decision rules, and the paths followed by the data reflect how the model makes predictions based on those features. To see the performance results of a model, you can look at the accuracy results on the training data and test data. The following are the accuracy results produced by the Decission Tree model.

```
Accuracy on Training Data: 1.00
Accuracy on Testing Data: 0.99


Classification Report on Testing Data:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        29
           1       1.00      0.97      0.98        31
           2       0.96      1.00      0.98        23

    accuracy                           0.99        83
   macro avg       0.99      0.99      0.99        83
weighted avg       0.99      0.99      0.99        83
```

Figure 6. Classification Report

Based on the model evaluation results shown in Figure 6, it can be seen that the accuracy on both data is very high, this shows that the model has very good performance with an accuracy level of 100% for the training data and 99% for the test data. This indicates that the model is able to classify eucalyptus leaves with high precision. However, 100% accuracy is not always good performance because it can be an indication of overfitting. So to ensure overfitting occurs or not, you need to consider several approaches, one of which is using a learning curve to see how well the model learns patterns from the data. This overfitting factor can be seen if the learning curve shows good performance on the training data but not on the validation data. Figure 7 is the result of the learning curve for this model.

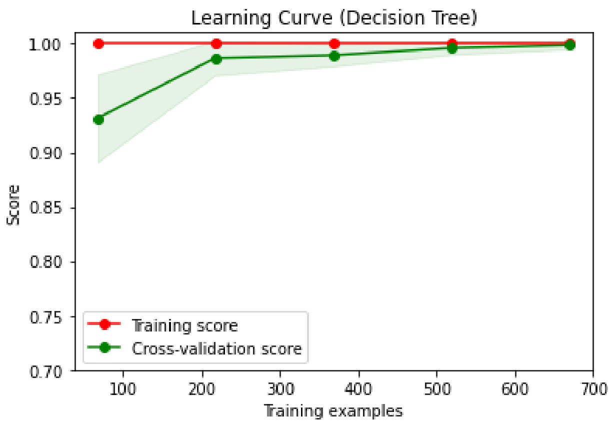

Figure 7. Learning Curve

Figure 6 shows that the model has good and consistent performance because the resulting curve is stable on training data and test data. This indicates that the model does not experience overfitting because the model can understand the training data well and can provide good predictions on new data. A stable curve on test data indicates that the model can generalize well on previously unseen data. Apart from using a learning curve, overfitting can also be ensured by performing cross validation and scoring the accuracy for each fold. The following are the accuracy results for each fold.

Table 1.
Accuracy of Each Fold in Cross Validation

K-Fold	Accuracy
Fold-1	0.9880
Fold-2	1.0000
Fold-3	0.9880
Fold-4	1.0000
Fold-5	10.000
Fold-6	0.9759
Fold-7	1.0000
Fold-8	1.0000
Fold-9	1.0000
Fold-10	1.0000
Average Accuracy	0.9952
Standard Deviation	0.0080

Table 1 shows that the model produces good performance at each fold, this is shown by the accuracy values obtained which are quite stable and do not depend too much on certain training data. This model produces an average accuracy value of 99.52% and a standard deviation of around 0.008, this indicates that the model performance is very stable and consistent in various experiments. The results of this evaluation provide confidence that the model used is an excellent choice for classifying the dataset used in this research.

4. Conclusion

The conclusion of this research shows that the use of Artificial Intelligence (AI) technology in combination with leaf image analysis has brought significant innovation in overcoming the challenges of identifying eucalyptus species in Indonesia. Through this approach, an effective and stable classification model was successfully developed using color histogram feature extraction and the Decision Tree algorithm. The evaluation results show a very high level of accuracy, with 100% for training data and 99% for test data. The validation process using learning curves and k-fold cross-validation confirmed that the model did not experience overfitting and had consistent performance. Thus, this study confirms that the proposed approach can be a reliable and efficient solution in automatically identifying eucalyptus species. This success has the potential to support better forest management and biodiversity conservation, as well as facilitating the development of a more responsible and environmentally friendly eucalyptus industry in Indonesia. By continuing to develop and apply AI technology in the forestry sector. This research development opens up opportunities for various developments and applications in various sectors. The automatic method developed to identify eucalyptus species from leaf images can be integrated in geographic information systems (GIS) for efficient forest mapping. This classification model can also be adapted to mobile applications for real-time species identification in the field. In the wood industry sector, this technology can improve wood management and trade by identifying resources correctly. Further development could expand the application of this technology to other plant species, increasing its relevance in the conservation and management of global biodiversity. By continuing to combine AI and image analysis, this research has the potential to become an important foundation in nature conservation and environmental sustainability

References

[1] O. C. Campoe *et al.*, “Climate and genotype influences on carbon fluxes and partitioning in Eucalyptus plantations,” *For. Ecol. Manage.*, vol. 475, no. May, 2020, doi: 10.1016/j.foreco.2020.118445.

[2] E. F. Elli, P. C. Sentelhas, and F. D. Bender, “Impacts and uncertainties of climate change projections on Eucalyptus plantations productivity across Brazil,” *For. Ecol. Manage.*, vol. 474, p. 118365, 2020, doi: https://doi.org/10.1016/j.foreco.2020.118365.

[3] S. M. G. Rocha *et al.*, “Influence of climatic variations on production, biomass and density of wood in eucalyptus clones of different species,” *For. Ecol. Manage.*, vol. 473, p. 118290, 2020, doi: https://doi.org/10.1016/j.foreco.2020.118290.

[4] D. P. dan P. K. Demak, “MENGENAL MANFAAT TANAMAN EUCALYPTUS,” *Dinas Pertanian dan Pangan Kabupaten Demak*, 2022. https://dinpertanpangan.demakkab.go.id/?p=4625 (accessed Feb. 14, 2024).

[5] M. Y. Salam, “Ekstraksi Minyak Atsiri dari Daun Eucalyptus grandis Menggunakan isolat Selulase,” *Skripsi*, vol. 2, pp. 1–89, 2018.

[6] T. A. Nasution, Yeni Herdiyeni, Wisnu Ananta Kusuma, Budi Tjahjono, and Iskandar Zulkarnaen Siregar, “Kecerdasan Buatan untuk Monitoring Hama dan Penyakit pada Tanaman Eucalyptus: Systematic Literature Review,” *J. Ilmu Komput. dan Agri-Informatika*, vol. 10, no. 2, pp. 224–237, 2023, doi: 10.29244/jika.10.2.224-237.

[7] V. B. Panunggul *et al.*, *PENGANTAR ILMU PERTANIAN*. Penerbit Widina, 2023.

[8] B. Dunia, *Pengembangan Pertanian Secara Berkelanjutan di Lahan*. Bank Dunia, 2021.

[9] L. Alonso, A. Rodríguez-Dorna, J. Picos, F. Costas, and J. Armesto, “Automatic differentiation of Eucalyptus species through Sentinel-2 images, Worldview-3 images and LiDAR data,” *ISPRS J. Photogramm. Remote Sens.*, vol. 207, pp. 264–281, 2024, doi: https://doi.org/10.1016/j.isprsjprs.2023.12.010.

[10] N. F. Amandari, I. Maulidina, and R. T. M. Akbar, “Studi Keanekaragaman dan Analisis Risiko Tumbuhan Invasif di Kawasan Taman Kelurahan Mekarmulya Kecamatan Cinambo, Kota Bandung, Provinsi Jawa Barat,” *J. Ris. Rumpun Ilmu Tanam.*, vol. 2, no. 2, pp. 51–61, 2023.

[11] J. Supriatna, *Konservasi Biodiversitas: Teori dan Praktik di Indonesia*. Yayasan Pustaka Obor Indonesia, 2018.

[12] N. Sutrisno and N. Heryani, “Teknologi konservasi tanah dan air untuk mencegah degradasi lahan pertanian berlereng,” *J. Penelit. dan Pengemb. Pertan.*, vol. 32, no. 3, 2014.

[13] A. Pramesta, “Perancangan Aplikasi Identifikasi Penyakit Kerdil dan Leaf Blight Pada Tanaman Lada Berdasarkan Citra Daun Berbasis Website,” Fakultas Sains dan Teknologi UIN Syarif HIdayatullah Jakarta.

[14] A. Muzammil, U. Islam, Z. Hasan, and G. Probolinggo, “Fakultas Tarbiyah Universitas Islam Zainul Hasan Genggong Probolinggo 88,” *Univ. Islam Zainul Hasan Genggong Probolinggo*, vol. 4, no. 2, pp. 88–112, 2022, [Online]. Available: file:///C:/Users/62831/Downloads/ahmad-muzammil-pba-1.pdf.

[15] M. Siahaan, C. H. Jasa, K. Anderson, M. V. Rosiana, S. Lim, and W. Yudianto, “Penerapan Artificial Intelligence (AI) Terhadap Seorang Penyandang Disabilitas Tunanetra,” *J. Inf. Syst. Technol.*, vol. 1, no. 2, pp. 186–193, 2020, [Online]. Available: https://journal.uib.ac.id/index.php/joint/article/view/4322.

[16] J. O. Ditinjau *et al.*, “The Ushuluddin International Student Conference,” vol. 1, no. 2, 2023, [Online]. Available: http://proceedings.radenfatah.ac.id/index.php/UInScof2022/manageIssues#future.

[17] F. A. Astuti, “Pemanfaatan Teknologi Artificial Intelligence untuk Penguatan Kesehatan dan Pemulihan Ekonomi Nasional,” *J. Sist. Cerdas*, vol. 4, no. 1, pp. 25–34, 2021, doi: 10.37396/jsc.v4i1.124.

[18] I. Gunawan, M. A. Liman, G. Ryan, and F. Purnomo, “Face Mask Detection for COVID-19 Prevention using Computer Vision,” *Procedia Comput. Sci.*, vol. 227, pp. 1143–1152, 2023, doi: https://doi.org/10.1016/j.procs.2023.10.626.

[19] A. K. Sharadhi, V. Gururaj, S. P. Shankar, M. S. Supriya, and N. S. Chogule, “Face mask recogniser using image processing and computer vision approach,” *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 67–73, 2022, doi: https://doi.org/10.1016/j.gltp.2022.04.016.

[20] A. A. Garcia Rubio, A. Garg, F. Yang, L. Li, B. Ozdoganlar, and P. R. LeDuc, “Image segmentation and control of freeform 3D ice printing with computer vision,” *Biophys. J.*, vol. 123, no. 3, Supplement 1, p. 548a, 2024, doi: https://doi.org/10.1016/j.bpj.2023.11.3318.

[21] A. Uryasheva, A. Kalashnikova, D. Shadrin, K. Evteeva, E. Moskovtsev, and N. Rodichenko, “Computer vision-based platform for apple leaves segmentation in field conditions to support digital phenotyping,” *Comput. Electron. Agric.*, vol. 201, p. 107269, 2022, doi: https://doi.org/10.1016/j.compag.2022.107269.

[22] J. Zhang, S. Qian, and C. Tan, “Automated bridge surface crack detection and segmentation using computer vision-based deep learning model,” *Eng. Appl. Artif. Intell.*, vol. 115, p. 105225, 2022, doi: https://doi.org/10.1016/j.engappai.2022.105225.

[23] B. Sreedha, P. R. Nair, and R. Maity, “Non-invasive early diagnosis of jaundice with computer vision,”

Procedia Comput. Sci., vol. 218, pp. 1321–1334, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.111.

[24] H. Li, N. Zeng, P. Wu, and K. Clawson, "Cov-Net: A computer-aided diagnosis method for recognizing COVID-19 from chest X-ray images via machine vision," *Expert Syst. Appl.*, vol. 207, p. 118029, 2022, doi: https://doi.org/10.1016/j.eswa.2022.118029.

[25] D. Đumić, D. Kečo, and Z. Mašetić, "Automatization of Microscopy Malaria Diagnosis Using Computer Vision and Random Forest Method," *IFAC-PapersOnLine*, vol. 55, no. 4, pp. 80–84, 2022, doi: https://doi.org/10.1016/j.ifacol.2022.06.013.

[26] J. Jaworek-Korjakowska, M. H. Yap, D. Bhattacharjee, P. Kleczek, A. Brodzicki, and M. Gorgon, "Chapter 3 - Deep neural networks and advanced computer vision algorithms in the early diagnosis of skin diseases," in *State of the Art in Neural Networks and Their Applications*, A. S. El-Baz and J. S. Suri, Eds. Academic Press, 2023, pp. 47–81.

[27] C. R. Pereira *et al.*, "A new computer vision-based approach to aid the diagnosis of Parkinson's disease," *Comput. Methods Programs Biomed.*, vol. 136, pp. 79–88, 2016, doi: https://doi.org/10.1016/j.cmpb.2016.08.005.

[28] M. Zortea *et al.*, "Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists," *Artif. Intell. Med.*, vol. 60, no. 1, pp. 13–26, 2014, doi: https://doi.org/10.1016/j.artmed.2013.11.006.

[29] S. Lee *et al.*, "Multimodal integration of neuroimaging and genetic data for the diagnosis of mood disorders based on computer vision models," *J. Psychiatr. Res.*, vol. 172, pp. 144–155, 2024, doi: https://doi.org/10.1016/j.jpsychires.2024.02.036.

[30] N. Hidayati, "EKSTRAKSI FITUR DENGAN COLOR HISTOGRAM DAN CLASSIFIER RANDOM FOREST PADA CITRA KUPU-KUPU," *JAMI J. Ahli Muda Indones.*, vol. 4, no. 2, pp. 42–51, 2023.

[31] R. Dijaya and H. Setiawan, "Buku Ajar Pengolahan Citra Digital," *Umsida Press*, pp. 1–85, 2023.

[32] D. Septhya *et al.*, "Implementasi Algoritma Decision Tree dan Support Vector Machine untuk Klasifikasi Penyakit Kanker Paru: Implementation of Decision Tree Algorithm and Support Vector Machine for Lung Cancer Classification," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 1, pp. 15–19, 2023.

[33] J. Im and J. R. Jensen, "A change detection model based on neighborhood correlation image analysis and decision tree classification," *Remote Sens. Environ.*, vol. 99, no. 3, pp. 326–340, 2005, doi: https://doi.org/10.1016/j.rse.2005.09.008.

[34] K. Topouzelis and A. Psyllos, "Oil spill feature selection and classification using decision tree forest on SAR image data," *ISPRS J. Photogramm. Remote Sens.*, vol. 68, pp. 135–143, 2012, doi: https://doi.org/10.1016/j.isprsjprs.2012.01.005.

[35] A. H. Rangkuti, Z. E. Rasjid, and D. J. Santoso, "Batik Image Classification Using Treeval and Treefit as Decision Tree Function in Optimizing Content Based Batik Image Retrieval," *Procedia Comput. Sci.*, vol. 59, pp. 577–583, 2015, doi: https://doi.org/10.1016/j.procs.2015.07.551.

[36] A. Jajodia *et al.*, "286P Improved diagnostic accuracy on MR imaging in post-surgical recurrent head and neck SCC lesions using decision tree classification system," *Ann. Oncol.*, vol. 31, p. S1353, 2020, doi: https://doi.org/10.1016/j.annonc.2020.10.280.

[37] C.-C. Yang *et al.*, "Application of decision tree technology for image classification using remote sensing data," *Agric. Syst.*, vol. 76, no. 3, pp. 1101–1117, 2003, doi: https://doi.org/10.1016/S0308-521X(02)00051-3.

[38] R. N. Ramadhon, A. Ogi, A. P. Agung, R. Putra, S. S. Febrihartina, and U. Firdaus, "Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank," *Karimah Tauhid*, vol. 3, no. 2, pp. 1860–1874, 2024.

[39] A. Lailiyah, V. R. S. Nastiti, E. D. Wahyuni, and C. S. K. Aditya, "Klasifikasi Tingkat Kemampuan Adaptasi Siswa dalam Pembelajaran Online Menggunakan Decision Tree," *Techno. Com*, vol. 23, no. 1, pp. 11–19, 2024.

[40] A. Y. Rahman and others, "Klasifikasi Citra Burung Lovebird Menggunakan Decision Tree dengan Empat Jenis Evaluasi," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 688–696, 2021.

[41] T. Novianti, S. A. Mandati, and E. K. Andana, "Peningkatan Evaluasi Risiko Kredit Menggunakan Decision Tree C 4.5," *MINE-TECH J. Manuf. Ind. Eng. \& Technol.*, vol. 2, no. 2, pp. 1–9, 2023.

[42] C. N. YUMNA, "Kaggle Repository," *Kaggle Repository*, 2023. https://www.kaggle.com/datasets/cutnasywayumna/eucalyptus-datasets.

[43] I. D. Id and U. Riau, "PERANCANGAN DAN PENERAPAN ALGORITMA STEMMING KATA Sebagai salah satu syarat untuk memperoleh gelar Sarjana dalam bidang Ilmu Komputer," no. March, 2018, doi: 10.13140/RG.2.2.16709.55526.

[44] V. Siahaan and R. H. Sianipar, *Panduan praktis dan komplet Pemrosesan Citra Digital dengan Matlab*. Balige Publishing, 2020.

[45] L. N. Rachmadi, A. P. Wibawa, and U. Pujianto, "Rekomendasi Jurusan dengan Menggunakan Decision Tree pada Sistem Penerimaan Peserta Didik Baru SMK Widya Dharma Turen," *Belantika Pendidik.*, vol. 4, no. 1, pp. 29–36, 2021.
[46] S. A. Pratiwi, A. Fauzi, S. A. P. Lestari, and Y. Cahyana, "Prediksi Persediaan Obat Pada Apotek Menggunakan Algoritma Decision Tree," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 4, pp. 2381–2388, 2024.
[47] D. Fatimah and E. Rahmawati, "Penggunaan Metode Decision Tree dalam Rancang Bangun Sistem Prediksi untuk Kelulusan Mahasiswa," *J. Algoritm.*, vol. 18, pp. 553–561, 2022, doi: 10.33364/algoritma/v.18-2.932.
[48] M. A. Wirya, "Deteksi Penyakit Alzheimer Pada Citra Magnetic Resonance Imaging Menggunakan Machine Learning dengan Metode Convolutional Neural Network," Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta.
[49] P. W. Rahayu *et al.*, *Buku Ajar Data Mining*. PT. Sonpedia Publishing Indonesia, 2024.
[50] D. Leni, "Pemilihan Algoritma Machine Learning Yang Optimal Untuk Prediksi Sifat Mekanik Aluminium," *J. Engine Energi, Manufaktur, dan Mater.*, vol. 7, no. 1, pp. 35–44, 2023.