Journal of Applied Science, Engineering, Technology, and Education *ISSN: 2685-0591 (online)*
Vol. 7 No. 1 (2025)
https://doi.org/10.35877/454RI.asci3772

Application of Digital Image Processing for Orchid Image Segmentation in Morphological Plant Analysis

Bob Subhan Riza, Rika Rosnelly, & Edy Victor Haryanto S.

Faculty of Engineering and Computer Science, Universitas Potensi Utama, Medan, Indonesia

Abstract

The research seeks to improve identification accuracy for orchid species by applying color-based segmentation to orchid images. The study utilized 90 images from the Kaggle database which represent Cattleya, Dendrobium, and Onchidium species. the first step Pre-processing the image as the initial procedure step requires resizing followed by color transformation from RGB to HSV space to prepare for future processing. The research team to do completed segmentation process with the K-Means method which clusters image pixels in accordance with color. The segmentation technique resulting with binary and RGB image outputs which accuracy 92% precision in difference flower objects from background image. The research project aims to the level accuracy better of segmentation process. The research demonstrates the application of digital image processing through K-Means segmentation to study orchid plant morphology. The method succeeded in enhancing species identification accuracy by extracting the floral items from their backgrounds across 90 photos of three orchid types. The findings support conservation work for endangered species while establishing strong foundations for developing orchid biodiversity research through enhanced image processing techniques. This research generates new possibilities for botanical and image processing technology applications.

Keywords: Digital Image Processing, Image Segmentation, Orchid, K-means

Received: 12 March 2025 Revised: 22 April 2025 Accepted: 27 April 2025

1. Introduction

This research using the K-Means method for segmentation orchid images. The segmentation method can excellent separation results with its good algorithm (Sabri *et al.*, 2016). Researchers already conduct in studies for species variation and ecological adaptation for the Orchidaceae family which reseacher can accurately differentiate with complex floral components. Selective can pressure from pollinator interactions processes can accelerates evolutionary transformations by enabling modularity for floral characteristics. (Artuso *et al.*, 2022). Furthermore, can integrating machine for some learning techniques (Andi Asrida Reskinah. D, Najib and Muhammad Ashdaq, 2025). Automated species for identification systems demonstrate which traditional procedure can with using machine learning algorithms boost for segmentation precision with image to produce reliable orchid biodiversity information (Zizka *et al.*, 2020). The author resulting exclusive orchid images with combines digital image with simple explanations. This study leads its discipline by merging with deep learning algorithms using color labels for reveal flower colors using a new method. For this research using five neural network are VGG16, Inception, Resnet50, Xception, and Nasnet underwent testing through different layer to unfreezing levels for determine the using optimal transfer learning configuration (Apriyanti *et al.*, 2021). using CNN method (Fajhar Muhammad, Agung Triayudi and Eri Mardiani, 2024) Methodology research of this which used segmentation generated image precise findings for analysis of the chosen sample of plant images. The research analysis all of images representations and decorative plant specimens are Black Orchid, Sirih Lurih, and Aglonema Triwarna (Lufila, Septyan Eka Prastya and Finki Dona Marleny, 2024). This research using CNN network for classification of plant leaf disease (Lu, Tan and Jiang, 2021).

Dataset the ORCHID provides a specialized database that enables researchers for further AI-based histopathological image analysis to oral cancer and early-stage oral cancer detection. Multiple centers contributed 300,000 image to patches to the ORCHID database which depict oral cancer and precancer lesions including oral squamous cell

* Corresponding author.

E-mail address: bob.potensi@gmail.com

carcinoma (OSCC), oral submucous fibrosis (OSMF). The dataset contains grade-level OSCC subclassifications including good differentiated (WD), moderately-differentiated (MD), and not good differentiated (PD) (Chaudhary *et al.*, 2023). This review paper discuss for multiple identification methods that aid in identification orchid species image to encourage sustainable conservation and combat invalid trading. This review explores upcoming using for these cutting edge with identification method (Nisruti, Naga Jogayya, Kothakota, Sheerin and Ampolu, 2020). A template matching algorithm can the research study classification orchid image by species and family using texture analysis. The testing step used data extracted image from a total of thirty images. The orchid image result a base correlation value of 0.7967 in relation to its correct label meaning remains higher than the classification result image's correlation value of 0.7899 (Andrio *et al.*, 2023). The study evaluation how Naive Bayes (NB)method and Tree Augmented Bayesian Network (TAN) method for feature classifiers in classification 63 orchid species. This Studies result which feature classifiers reach accuracy levels between 83% and 93% (Apriyanti, Spreeuwers and Lucas, 2023). This study developed machine learning methode for leaf area prediction models for Dendrobium nobile using multiple linear regression (MLR), support vector regression (SVR), gradient boosting regression (GBR), ANNs (Das *et al.*, 2022). The research using the Naive Bayes method for the classification of texture features extracted image from black orchid leaves. The testing results showed ladybirds at 1.18 snails scored at 1.43 mites at 1.27 and caterpillars at 1.22.

The confusion matrix generated ideal results by achieving complete accuracy 100%, precision 100% and recall at 100% (Ismayanti *et al.*, 2022). This study approach employs the Xception model with one untrained layer for feature extraction to classify six unique orchid traits. The classifier is composed of two thick layers separated by 0.5 dropouts. The completed model achieved a macro-averaged f1 score of 0.85 (Post, 2020). This approach seeks to classify orchid plants based on their specific types. Researchers used sets of photos numbering over 2250 for training and 1500 for testing across 15 distinct categories. The research team implemented SVM with linear, polynomial, and Gaussian kernels while k-nearest neighbor operated using distances between K1 and K11. Naïve Bayes achieved 66% accuracy which is lower than SVM but the highest KNN which reached 98% accuracy used K=1 and d=1 (Andono *et al.*, 2021). Because of the image data limitation for deep CNN-based classification of rare orchids from Indonesia the researcher introduced a multi-stage image data augmentation method as a solution. A ResNet CNN model which utilizes transfer learning techniques is implemented for classification purposes. Experiments with multiple data augmentation techniques as well as stage 1 and stage 2 augmentation datasets show that the proposed system delivers superior performance versus existing methods (Dewantara *et al.*, 2020) A classification system for orchids operates through Convolutional Neural Network technology (Kattenborn *et al.*, 2021; Ideastari *et al.*, 2024), IoT (Chen *et al.*, 2022). A new multi-technique method enabled researchers to identify five wild orchid species which include common spotted-orchid (Dactylorhiza fuchsia), heath spotted-orchid (Dactylorhiza maculata), pyramidal orchid (Anacamptis pyramidalis), heath fragrant-orchid (Gymnadenia borealis), and dark-red helleborine (Epipactis atrorubens) (Ahmed *et al.*, 2024), Cattleya trianae (Forero *et al.*, 2020), Plant Leaf Diseases(Tugrul, Elfatimi and Eryigit, 2022; Khandelwal *et al.*, 2024), Convolvulus sepium detection (Gao *et al.*, 2020), Machine learning and Deep learning(Barhate *et al.*, 2024), (Lu *et al.*, 2022), (Haryanto *et al.*, 2017),(Singh *et al.*, 2016),(Sarma, Boruah and Buragohain, 2023), AI (Shaikh, Rasool and Mir, 2025), (Ayoub Shaikh, Rasool and Rasheed Lone, 2022).

2. Methods

This research employed 90 total images composed of 30 images for each of three orchid types: cattleya, dendrobium, and onchidium which were presented in jpg format.

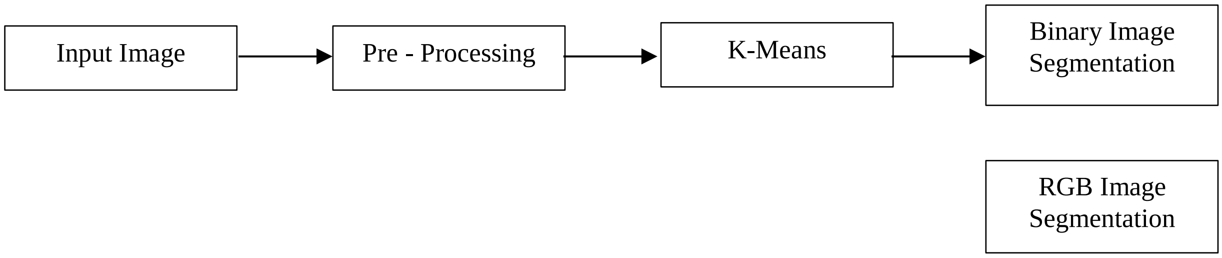

Fig. 1. Block Diagram Processing

3. Results and Discussion

The outcome of the study approach is the model that was created using it, which incorporates a conversation at all stages**.**

3.1 Input Image

Figure 2 describes one type of orchid used in this study and the data in this study were taken from kaggle.com as much as 90 data, each type consisting of 30 images.

Fig. 2. Cattleya

3.2 Preprocessing

The pre-processing stage is that the image is reduced in size to 25% of the original size to speed up computation, RGB image is extracted into three colour components (R, G, B) for further analysis, each colour component is adjusted in contrast to improve the visual quality of the image, RGB image is converted into HSV colour space, which is more suitable for segmentation based on colour and the results of pre-processing can be seen in Figure 3.

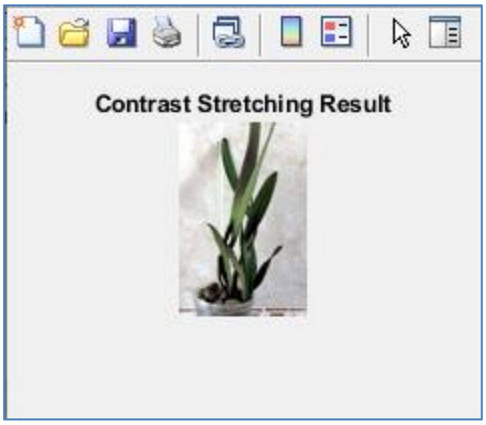

Fig. 3. Contrast Stretching

3.3 Segmentation using K-Means

The RGB image is converted to HSV (Hue, Saturation, Value) colour space for further analysis, then the results of the colour transformation are extracted features performed on the H and S components of the HSV image, then a segmentation stage is performed using the K-Means algorithm which serves to group pixels based on the extracted features, existing pixels will be labelled based on the clusters generated by the K-Means algorithm and the cluster with the smallest area is identified as the existing cluster and the results of pre-processing can be seen in Figure 4.

Each pixel x_i is grouped into the cluster j that minimises the squared distance to the centroid of that cluster. A commonly used distance function is the Euclidean distance, which is expressed by the formula:

$$j = \arg\min_k \|x_i - C_k\|^2$$

where:

x_i is the feature representation of pixel
C_k is the centroid represents the central point of the k cluster.
$\| x_i - C_k \|^2$ the centroid represents the central point of the k cluster.

$$\|x_i - C_k\|^2 = (x_i[1] - C_k[1])^2 + (x_i[2] - C_k[2])^2 + \ldots + (x_i[m] - C_k[m])^2$$

After assigning labels to all pixels the k cluster centroid becomes the average position of its member pixels. The formula is:

$$C_k = \frac{1}{N_k} \sum_{x_i \in Cluster_k} x_i$$

N_k is the number of pixels belonging to the k cluster.
$x_i \in$ *Cluster k* are all the pixels labelled to the k cluster.

The clustering and centroid updating routine continues until convergence occurs. The algorithm reaches convergence when the centroid stops showing significant changes or remains stable in cluster assignments.

The convergence criteria can be expressed as:

$$\|C_k^{(t)} - C_k^{(t-1)}\| < \epsilon$$

ϵ is a small threshold value that determines when the process can be stopped.

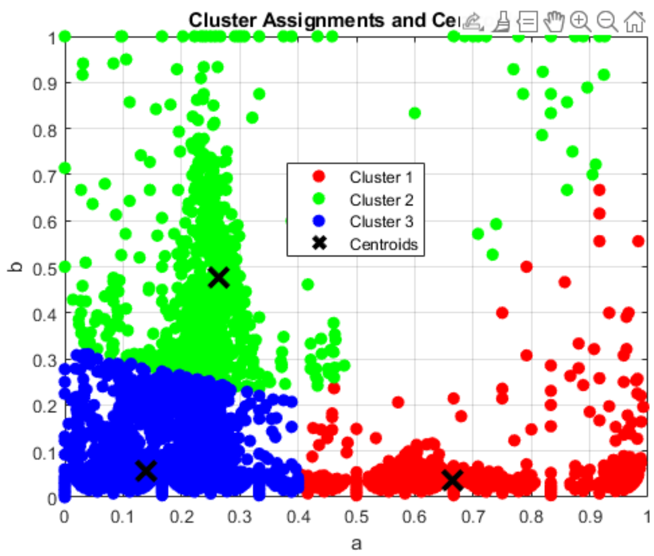

Fig. 4. Result centroid by K-Means

The study's segmentation results demonstrate how well K-Means segmentation separates orchid blooms from their backgrounds. However, different orchid species have varying segmentation performance. While species like

Onchidium, which may have smaller and more complex blooms, can provide difficulties due to minor color changes and overlapping petal structures, species like Cattleya, which have huge, distinct petals, may present less uncertainty during segmentation.

3.4 Binary Image Segmentation

The binary segmentation outcome depicted in Figure 5 represents an image with just two possible pixel values, either 0 or 1. Pixels assigned a value of 1 represent the object of flower while pixels assigned a value of 0 represent background regions. Numerical analysis and subsequent processing benefit greatly from this outcome. The thresholding or clustering process produces an image where orchid flower pixels receive a value of 1 but all other pixels receive the value 0. The segmentation results benefit from morphological operations including erosion and dilation which also serve to reduce noise and fill object gaps. Machine learning algorithms and pattern recognition systems can process binary images as their input data.

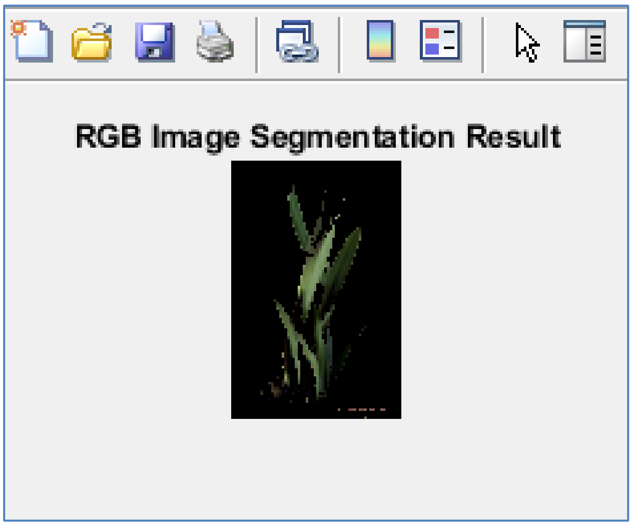

Fig. 5. Result RGB Image Segmentation

A number strategies can be used to alleviate the problems caused by image quality, such as illumination and resolution:

a. **Data Augmentation**: Methods like as resizing, flipping, and rotating photos, along with modifying contrast and brightness, can assist replicate different environmental conditions and increase the model's resilience.
b. **Pre-processing Techniques**: To improve image contrast and reduce lighting irregularities, image enhancement techniques such as histogram equalization or CLAHE (Contrast Limited Adaptive Histogram Equalization) can be used.
c. **Multi-scale Segmentation**: Flowers with complex shapes or structures benefit greatly from the ability to capture both fine details and broader features through the use of segmentation at various scales

4. Conclusion

The segmentation method is greatly impacted by the quality of the orchid photos, particularly by environmental elements like lighting and image resolution. Inconsistencies in the image's color distribution might be introduced by lighting variations such shadows, overexposure, or uneven illumination. This discrepancy can make it more difficult to separate the orchid blossoms from the background, which is particularly important when using techniques that mostly rely on color features, such as K-Means clustering. Furthermore, dim lighting might mask important flower features, making precise segmentation more difficult.

The research demonstrates the application of digital image processing through K-Means segmentation to study orchid plant morphology. The method succeeded in enhancing species identification accuracy by extracting the floral items from their backgrounds across 90 photos of three orchid types. The combined pre-processing procedures of image resizing and color component extraction with HSV conversion and RGB binary segmentation resulted in visualizations useful for further research.

The findings support conservation work for endangered species while establishing strong foundations for developing orchid biodiversity research through enhanced image processing techniques. This research generates new possibilities for botanical and image processing technology applications.

The study's segmentation strategy works well, but for more precise plant picture categorization, it might profit from integrating current developments in machine learning and deep learning. In recent years, a number of studies have concentrated on the identification and segmentation of plant species using convolutional neural networks (CNNs) and deep learning models. Recent studies, for example, have effectively used deep learning models such as VGG16, ResNet, and Xception for classification tasks, demonstrating greater accuracy than conventional techniques. These methods could improve the model's capacity to identify intricate floral patterns and manage species-specific variations more accurately. These techniques usually entail transfer learning, which uses large, varied datasets to improve generalization by fine-tuning previously taught models on orchid datasets.

It would be helpful to provide thorough explanations of how the algorithm handles various species in order to enhance comprehension of segmentation findings across species. For example, while Cattleya has greater floral components, the segmentation process may yield clear, well-defined segments, yet Dendrobium's intricate flower structure may result in more fragmented segments. The approach's advantages and disadvantages might be better illustrated by a qualitative comparison of the segmentation outputs, maybe using side-by-side visual comparisons. Furthermore, adding quantitative performance indicators for each species, including precision, recall, and F1-score, would give a better idea of how well the model generalizes to other orchid species.

References

Ahmed, S. *et al.* (2024) 'Use of remote sensing and image processing for identification of wild orchids', *Frontiers in Environmental Science*, 12. Available at: https://doi.org/10.3389/fenvs.2024.1371445.

Andi Asrida Reskinah. D, Najib, M. and Muhammad Ashdaq (2025) 'Sentiment Analysis on Erspo Jersey in X Using Machine Learning Algorithms', *SAGA: Journal of Technology and Information System*, 2(3), pp. 288–294. Available at: https://doi.org/10.58905/saga.v2i3.334.

Andono, P. *et al.* (2021) 'Orchid types classification using supervised learning algorithm based on feature and color extraction', *Bulletin of Electrical Engineering and Informatics*, 10, pp. 2530–2538. Available at: https://doi.org/10.11591/eei.v10i5.3118.

Andrio, A. *et al.* (2023) 'Application Of The Template Matching Method In Orchid Image Classification In Bengkulu', *Jurnal Komputer, Informasi dan Teknologi*, 3(1), pp. 151–158. Available at: https://doi.org/10.53697/jkomitek.v3i1.1197.

Apriyanti, D., Spreeuwers, L. and Lucas, P.J. (2023) 'Deep neural networks for explainable feature extraction in orchid identification', *Applied Intelligence*, 53, pp. 1–16. Available at: https://doi.org/10.1007/s10489-023-04880-2.

Apriyanti, D.H. *et al.* (2021) 'Automated color detection in orchids using color labels and deep learning', *PLoS ONE*, 16(10 October), pp. 1–27. Available at: https://doi.org/10.1371/journal.pone.0259036.

Artuso, S. *et al.* (2022) 'Evidence for an evo-devo-derived hypothesis on three-dimensional flower shape modularity in a tropical orchid clade', *Evolution*, 76(11), pp. 2587–2604. Available at: https://doi.org/10.1111/evo.14621.

Ayoub Shaikh, T., Rasool, T. and Rasheed Lone, F. (2022) 'Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming', *Computers and Electronics in Agriculture*, 198, p. 107119. Available at: https://doi.org/10.1016/J.COMPAG.2022.107119.

Barhate, D. *et al.* (2024) 'A systematic review of machine learning and deep learning approaches in plant species detection', *Smart Agricultural Technology*, 9, p. 100605. Available at: https://doi.org/10.1016/J.ATECH.2024.100605.

Chaudhary, N. *et al.* (2023) 'ORCHID: A Comprehensive Oral Cancer Histology Image Database for Histopathological Analytics and Diagnostics', *medRxiv*, 11(1), pp. 1–14.

Chen, L.-B. *et al.* (2022) 'A Self-Supervised Learning-Based Intelligent Greenhouse Orchid Growth Inspection System for Precision Agriculture', *IEEE Sensors Journal*, 22, pp. 24567–24577. Available at: https://doi.org/10.1109/JSEN.2022.3221960.

Das, M. *et al.* (2022) 'A Machine Learning Approach for the Non-Destructive Estimation of Leaf Area in Medicinal Orchid Dendrobium nobile L.', *Applied Sciences*, 12(9). Available at: https://doi.org/10.3390/app12094770.

Dewantara, D.S. *et al.* (2020) *CNN with Multi Stage Image Data Augmentation Methods for Indonesia Rare and Protected Orchids Classification*. Available at: https://doi.org/10.1109/ICOSICA49951.2020.9243174.

Fajhar Muhammad, Agung Triayudi and Eri Mardiani (2024) 'Implementation of Face Recognition for Lecturer Attendance Using Deep Learning CNN Algorithm', *SaNa: Journal of Blockchain, NFTs and Metaverse Technology*, 2(2), pp. 123–130. Available at: https://doi.org/10.58905/sana.v2i2.275.

Forero, M. *et al.* (2020) 'Classification of Cattleya Trianae and Its Varieties by Using Colorimetry', in, pp. 35–44. Available at: https://doi.org/10.1007/978-3-030-49076-8_4.

Gao, J. *et al.* (2020) 'Deep convolutional neural networks for image-based Convolvulus sepium detection in sugar beet fields', *Plant Methods*, 16. Available at: https://doi.org/10.1186/s13007-020-00570-z.

Haryanto, S.E.V. *et al.* (2017) 'A fast and accurate detection of Schizont plasmodium falciparum using channel color space segmentation method', in *2017 5th International Conference on Cyber and IT Service Management, CITSM 2017*. Institute of Electrical and Electronics Engineers Inc. Available at: https://doi.org/10.1109/CITSM.2017.8089290.

Ideastari, N. *et al.* (2024) 'An optimum hyperparameters of restnet-50 for orchid classification based on convolutional neural network', *Journal of Soft Computing Exploration*, 5, pp. 55–66. Available at: https://doi.org/10.52465/joscex.v5i1.297.

Ismayanti, R. *et al.* (2022) *Identification of Pests on Black Orchid Plants Using Naïve Bayes Method Based on Leaf Image Texture*. Available at: https://doi.org/10.1109/ICSINTESA56431.2022.10041471.

Kattenborn, T. *et al.* (2021) 'Review on Convolutional Neural Networks (CNN) in Vegetation Remote Sensing', *ISPRS Journal of Photogrammetry and Remote Sensing*, 173, pp. 24–49. Available at: https://doi.org/10.1016/j.isprsjprs.2020.12.010.

Khandelwal, S. *et al.* (2024) 'Optimizing Performance in Mango Plant Leaf Disease Classification through Advanced Machine Learning Techniques', *Engineering, Technology & Applied Science Research*, 14(6), pp. 18476–18480. Available at: https://doi.org/10.48084/etasr.8220.

Lu, J., Tan, L. and Jiang, H. (2021) 'Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification', *Agriculture*, 11(8). Available at: https://doi.org/10.3390/agriculture11080707.

Lu, Y. *et al.* (2022) 'Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review', *Computers and Electronics in Agriculture*, 200, p. 107208. Available at: https://doi.org/10.1016/J.COMPAG.2022.107208.

Lufila, L., Septyan Eka Prastya and Finki Dona Marleny (2024) 'Image Segmentation Of Ornamental Plants Typical Of South Kalimantan Using The Convolutional Neural Network Method', *Install: Information System and Technology Journal*, 1(1), pp. 15–22. Available at: https://doi.org/10.33859/install.v1i1.544.

Nisruti, A., Naga Jogayya, Kothakota , Sheerin, B. and Ampolu, S. (2020) 'Identification Of Orchidaceae Species by Using Various Methods: A Review', *Journal of Emerging Technologies and Innovative Research (JETIR)*, 06(4), pp. 7–20.

Post, C.A. (2020) 'Multilabel Classification of Orchid Features based on Deep Learning', pp. 1–10.

Sabri, B. *et al.* (2016) *K-Means vs. Fuzzy C-Means for Segmentation of Orchid Flowers*. Available at: https://doi.org/10.1109/ICSGRC.2016.7813306.

Sarma, P., Boruah, P.A. and Buragohain, R. (2023) 'MED 117: A dataset of medicinal plants mostly found in Assam with their leaf images, segmented leaf frames and name table', *Data in Brief*, 47, p. 108983. Available at: https://doi.org/10.1016/J.DIB.2023.108983.

Shaikh, T.A., Rasool, T. and Mir, W.A. (2025) 'Fields of the future: Digital transformation in smart agriculture with large language models and generative AI', *Computer Standards & Interfaces*, 94, p. 104005. Available at: https://doi.org/10.1016/J.CSI.2025.104005.

Singh, A. *et al.* (2016) 'Machine Learning for High-Throughput Stress Phenotyping in Plants', *Trends in Plant Science*, 21(2), pp. 110–124. Available at: https://doi.org/10.1016/J.TPLANTS.2015.10.015.

Tugrul, B., Elfatimi, E. and Eryigit, R. (2022) 'Convolutional Neural Networks in Detection of Plant Leaf Diseases: A Review', *Agriculture*, 12(8). Available at: https://doi.org/10.3390/agriculture12081192.

Zizka, A. *et al.* (2020) 'Automated conservation assessment of the orchid family with deep learning', *Conservation Biology*, 35. Available at: https://doi.org/10.1111/cobi.13616.