iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83435

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Custom Convolutional Neural Networks and Automated Machine Learning for Medicinal Plant Identification: A Comparative Review

Syeda Asra Arshia Qadri[1]
[1]*Assistant Professor, Department of Computer Science, Dr. Rafiq Zakaria College for Women, Aurangabad, Maharashtra, India*

Abstract: ***The botanical misidentification of medicinal herbs, particularly the accidental substitution of therapeutic flora with toxic look-alikes, presents critical risks to global public health and pharmaceutical safety. This review synthesizes technical advancements and performance benchmarks across ten target papers, contrasting hand-tuned pipelines—including attention-enhanced EfficientNetV2, Swin-Transformers, and multi-granularity architectures—against emerging Automated Machine Learning (AutoML) frameworks.***
While stochastic ensembles and custom ResNet variants achieve accuracies exceeding 95%, substantial classification fragility persists when models transition from pristine laboratory imagery to unstructured wild environments. Primary dataset bottlenecks include signal interference from dynamic water reflections, soil-cluttered backgrounds, and the failure of networks to resolve intricate leaf vein geometry in dense, overlapping growth.
The strategic technological outlook emphasizes the imperative for lightweight, parameter-efficient architectures suited for edge deployment, coupled with Explainable AI (XAI) tools like Grad-CAM.
Such frameworks are essential to mitigate the "double black box" of algorithmic opacity and traditional medical theory, providing field botanists with transparent diagnostic reasoning.
Keywords: Medicinal Plant Recognition; Convolutional Neural Networks; Automated Machine Learning (AutoML); Fine-grained Image Classification; Explainable AI (XAI).

I. INTRODUCTION

The global significance of medicinal plants is established by their role as foundational resources for Traditional Chinese Medicine (TCM) and modern pharmaceutical discovery. These botanical taxa are frequently employed to mitigate complex clinical challenges, including malignant neoplasms and infectious disease outbreaks. However, the clinical efficacy of such treatments is entirely dependent upon rigorous botanical authentication. Severe pharmaceutical risks are introduced when medicinal flora are misidentified or substituted with toxic look-alikes. Fatal outcomes, acute allergic reactions, and systemic organ failure are documented consequences of the accidental ingestion of toxic species. The presence of taxonomically close mimics, such as species within the *Acorus* genus, presents significant character similarities that complicate identification and compromise patient safety.

Botanical authentication processes were historically reliant on manual botanical inspection, which is characterized as a laborious, subjective, and error-prone endeavor. A digital transformation is currently being realized through the integration of computer vision sensors and deep learning architectures into identification pipelines. The field has shifted from the use of handcrafted feature descriptors to the implementation of Convolutional Neural Networks (CNNs), which enable the automated extraction of discriminative hierarchical features. While custom-tuned architectures like ResNet, VGG, and Vision Transformers remain central to the domain, an operational transition toward accessible Automated Machine Learning (AutoML) systems is identified. These automated frameworks are utilized to minimize the time-intensive overhead associated with manual hyperparameter optimization and data labeling.

The objective of this narrative review is the critical evaluation of architectural performance and dataset vulnerabilities across 10 foundational research papers. Technical constraints related to data scarcity and classification fragility are examined, particularly the failure of networks to generalize when transitioned from laboratory settings to unstructured wild environments. This review is structured as follows: we synthesizes the parameters of the reviewed literature; contrasts the performance of custom CNNs against AutoML tools; analyzes morphological mimicry and ecological noise; and identifies critical gaps and the requirement for explainable AI.

©IJRASET: All Rights are Reserved

Reference Marker	Botanical Target Species	Algorithmic Framework Implemented	Primary Metrics	Critical Technical Constraints
[Sendhilkumar, 2024]	2626 unique Indian medicinal plant species	Stochastic weight averaging ensemble of ResNet50, MobileNetV2, and VGG16	95.26% Overall Accuracy	Failure to recognize dry plant parts; significant challenges in comprehensive dataset creation
[Zou, 2025]	TCMP-300 benchmark (300 categories) and 7 representative herbs (e.g., *Panax ginseng*)	Swin-Transformer backbone with Bidirectional Semantic Transformer and Gradient Optimization Module (GOM)	90.32% Accuracy (TCMP-300); 92.75% Accuracy (Self-built)	High intra-class morphological similarity; complex backgrounds and varying imaging conditions
[Miao, 2023]	Six categories including *Angelica sinensis* radix and *Citrus reticulatae*	Improved ConvNeXt (fused with ACMix and Stacked FFN)	85.2% verification accuracy (22-layer depth)	Scarcity of publicly available labeled image samples; sensitivity to environmental signal noise
[Gamage, 2024]	15 varieties of Sri Lankan Ayurvedic plants	Pre-trained VGG16 CNN for identification; PyCaret AutoML for remedy classification	96.88% Accuracy (VGG16); 93.33% Accuracy (KNN remedy model)	Insufficiency of training data; language barriers (Sinhala vs. English); noisy web-scraped images
[Zhu, 2026]	47 categories of toxic Chinese medicinal herbs	Attention-enhanced EfficientNetV2 (Multi-Scale Feature Fusion + CBAM)	91.28% Top-1 Accuracy; 90.27% Macro F1-score	Real-world background clutter; small and indistinct target regions; subtle inter-class visual similarities
[Guan, 2024]	*Brasenia schreberi* (medicinal aquatic vegetable)	YOLO-GS (lightweight Ghost convolution + C3-GS cross-stage module)	95.7% Average Accuracy	Limited computational power of industrial control computers; water surface reflections; overlapping leaf occlusion
[Zhao, 2024]	*Astragalus membranaceus* (Huangqi) slices	FA-SD-YOLO (improved YOLOv8n with AIFI module and SD detection head)	88.6% Accuracy; 89.1% F1-score; 93.2% mAP	High morphological similarity between healthy and defective slices; misclassification of "special-shaped" targets
[Wu, 2026]	Comprehensive TCM herb categories	Survey including LLMs (TCM-KLLaMA, TCM-FTP) and deep learning architectures	0.8031 F1-score (TCM-FTP prescription prediction)	"Double black boxes" (opacity of algorithm and theory); lack of unified naming/data standards
[Liu, 2025]	*Acorus tatarinowii* (Shi Chang Pu) and *Acorus calamus* (Shui Chang Pu)	ResNet50-SE (channel attention) and YOLOv8 with spatial attention	94.5% Accuracy (ResNet50-SE); 98.6% Accuracy (YOLOv8)	Regional sample bias (Southern China focus); high GPU computational resource demands
[Yang, 2026]	100 species of medicinal plants (MP-100 dataset)	MGCA-Net (ResNet-50 backbone + Jigsaw Patch Generator + Cross-Attention)	92.40% Top-1 Accuracy	Insufficient samples for rare/endangered taxa; performance degradation in low-light or cluttered backgrounds

Table 1: Parameters Of The Reviewed Literature

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

II. ARCHITECTURAL EVALUATION: CUSTOM CNNS VS. AUTOML

The architectural landscape of botanical recognition in the reviewed literature is defined by a primary dichotomy between hand-tuned deep learning pipelines and emerging automated machine learning (AutoML) frameworks. Structural performance is largely dictated by the ability of an architecture to handle high intra-class morphological similarity and complex environmental noise. Hand-tuned models, such as the ResNet and VGG families, remain the dominant benchmarks. A stochastic weight averaging ensemble of ResNet50, MobileNetV2, and VGG16 was demonstrated to reach an overall accuracy of 95.26%, leveraging the diverse learning patterns of independent architectures to mitigate individual biases.[1]

Furthermore, VGG16 was identified as the superior individual classifier for Ayurvedic species, achieving 96.88% accuracy through frozen convolutional bases and bottleneck feature refinement [2].

Advanced structural modifications are frequently implemented in custom pipelines to address domain-specific challenges like background clutter. The MGCA-Net architecture utilizes a ResNet-50 backbone augmented by a Jigsaw Patch Generator and multi-scale progressive learning to achieve 92.40% Top-1 accuracy [8]. Similarly, the integration of Squeeze-and-Excitation (SE) modules into ResNet50 and spatial attention into YOLOv8 has been shown to enhance classification and object detection by 1.7% and 1.2% respectively [6]. Transformer-based models, specifically the Swin-Transformer, exhibit robust performance (90.32% accuracy) when combined with bidirectional semantic modeling and graph attention networks [5]. However, these architectures demonstrate significant sensitivity to dataset scale; without extensive pre-training, Vision Transformers (ViT) may suffer from diffuse attention, resulting in accuracies as low as 30.99% compared to CNN-based baselines [10].

The hyperparameter optimization (HPO) overhead represents a critical differentiator between these paradigms. Hand-tuned pipelines typically necessitate labor-intensive grid searches or random searches for learning rates, weight decays, and batch sizes [5], [10]. In contrast, AutoML frameworks like PyCaret implement automated feature selection and HPO through functions such as tune_model(), which utilizes Random Grid Search to optimize hyperparameters across a predefined space [2]. Such platforms are recognized for reducing the "laborious and time-consuming" nature of manual configuration, facilitating the identification of upwards of 315 herbal categories [7].

Ease of deployment is primarily evaluated through computational efficiency and parameter reduction. For unmanned harvesting and edge computing, lightweight hand-tuned models are prioritized. The YOLO-GS model implements Ghost convolutions and a C3-GS cross-stage module to reduce parameter counts by 46.5%, achieving a computational cost of only 9.5 GFLOPS [9]. Similarly, the FA-SD-YOLO model reduces parameters by 47.3% through the integration of FasterNet modules and a shared convolutional detection head [9]. While AutoML offers a simplified development cycle, hand-tuned models currently provide more granular control over deployment metrics on resource-constrained industrial control computers [9]. However, both paradigms face the challenge of "double black boxes"—the opacity of algorithmic logic combined with the complexity of traditional medical theory [7].

III. DATASET BOTTLENECKS AND ECOLOGICAL VARIATIONS

The inherent dichotomy between laboratory-acquired imagery and unstructured wild environments introduces a critical fragility in medicinal plant classification. Investigations reveal that models trained on pristine datasets—characterized by clean backgrounds, uniform lighting, and centrally positioned objects—demonstrate a pronounced inability to generalize in practical field scenarios. Real-world field noise, consisting of complex soil backgrounds, dynamic lighting variations, and unpredictable partial occlusions, substantially lowers the signal-to-noise ratio and impairs hierarchical feature extraction. For instance, in unmanned aquatic harvesting, the unique interference of water surface reflections triggered by sunlight changes acts as a pervasive interfering factor that obscures botanical target details. Furthermore, the lack of generalizability in existing public datasets, which often rely on isolated single-leaf samples, fails to account for the "intensity signals" and environmental noise that disrupt pixel correlations in native habitats. Consequently, the transition from controlled laboratory conditions to unconstrained field environments results in "leakage detection" and misclassification, as the networks struggle to differentiate botanical targets from the surrounding visual clutter.

Beyond environmental noise, morphological mimicry and phenotypic plasticity represent lethal failure modes for automated systems, particularly when distinguishing between therapeutic flora and toxic look-alikes. The challenge of fine-grained recognition is intensified by subtle inter-class similarities in color, texture, and morphology that demand "instance-level" sensitivity. Differentiating between true medicinal species and their toxic counterparts requires the isolation of micro-texture cues, such as precise leaf vein geometry and petal venation. However, authors document that "overlapping occlusion"—caused by dense growth where leaves crisscross and intermix—leads to a fundamental loss of these discriminative local features.

©IJRASET: All Rights are Reserved

This failure mode is especially prevalent in taxonomically close species, where the "diffuse attention" of networks fails to capture the intricate geometric patterns of leaf veins necessary for separation. For example, the high degree of similarity between "special-shaped" and "special-shaped_rot" *Astragalus* targets frequently complicates identification, leading to misdiagnosis. Similarly, species like *Acorus tatarinowii* and *Acorus calamus* exhibit "character similarities" that pose significant risks to clinical safety when their distinct phenotypic manifestations are obscured by overlapping spatial arrangements. This "double opacity"—the lack of algorithmic interpretability combined with the complexity of botanical theory—means that without robust multi-granularity modeling, networks cannot reliably isolate the specific leaf vein geometry required to prevent the accidental ingestion of toxic mimics. Ultimately, these failure modes underscore a critical bottleneck: the reliance on statistical regularities rather than genuine structural reasoning leads to a "black box" logic that is insufficient for high-risk clinical or toxicological applications.

IV. CRITICAL GAPS AND OPEN RESEARCH CHALLENGES

The synthesis of current literature reveals that the most significant bottleneck in medicinal plant recognition is the global deficit of centralized, open-access, and botanically verified image databases. Multiple authors document a pervasive "lack of publicly available datasets" for traditional herbal medicine, forcing researchers to rely on Bing-scraped images or small, self-constructed repositories that lack scale. These fragmented data sources often fail to represent real-world variability, as most existing benchmarks are composed of single-leaf samples captured under idealized laboratory conditions. This "idealized" training prevents models from generalizing to field environments characterized by complex backgrounds, water reflections, and partial occlusions. Furthermore, the literature identifies a critical shortage of data for rare or endangered taxa and high-risk subdomains like toxic herbs, where the consequences of misclassification are medically severe. The lack of unified naming standards and regional sample biases—such as focus restricted to Southern China—further complicates the development of a universal botanical recognition infrastructure.

Simultaneously, there is an urgent technical imperative for Explainable AI (XAI) frameworks to mitigate the "double black box" problem—the inherent opacity of deep learning algorithms combined with the complexity of traditional medical theory. High-performing models currently operate as opaque predictors, providing accurate results without revealing the underlying decision logic. This lack of transparency is unacceptable for professional botanists and pharmacists who require transparent reasoning pathways to validate therapeutic recommendations. Authors have begun addressing this through Gradient-weighted Class Activation Mapping (Grad-CAM), which generates weighted localization maps to visualize whether a network is focusing on botanically relevant regions, such as the junction between a fungal stem and insect body or specific leaf vein geometry. Such XAI tools are documented to provide "intuitive evidence" of a model's structural awareness, allowing experts to confirm that "machine experience" aligns with established pharmacognosy. Visualization of intermediate feature maps demonstrates that these frameworks can effectively highlight class-discriminative features while suppressing irrelevant environmental noise. Ultimately, transitioning AI from opaque black boxes to interpretable diagnostic aids is essential for fostering clinical trust and securing regulatory acceptance in the pharmaceutical sector.

V. FUTURE PERSPECTIVES AND STRATEGIC ROADMAP

The progression of botanical recognition research necessitates a strategic transition from centralized, cloud-dependent computing toward decentralized Mobile Edge AI deployment. Current deep learning pipelines, while accurate, often exhibit high computational overhead that precludes execution on resource-constrained hardware in remote field environments [1] [9]. To empower field botanists and indigenous communities with offline identification capabilities, future research must prioritize the development of quantized AutoML models. Techniques such as knowledge distillation and the implementation of lightweight convolutional modules—exemplified by Ghost-convolutions and depth-wise separable filters—must be refined to enable real-time inference on edge devices without compromising Top-1 accuracy [9], [6]. Such localized deployment ensures that botanical authentication can be performed in native habitats where internet connectivity is frequently unavailable. [1] [7]

Furthermore, to resolve the critical failure modes associated with phenotypic plasticity and morphological mimics, the adoption of multimodal data fusion is essential. The inherent limitations of 2D leaf geometry in distinguishing taxonomically close therapeutic flora from toxic counterparts require the integration of complementary analytical streams [9], [10]. Future frameworks should fuse computer vision features with chemical spectroscopy profiles (e.g., Raman or metabolomics) and localized genomic data, such as DNA barcoding sequences [5] [7] [6]. This multi-dimensional evidence base will allow for "instance-level" sensitivity, effectively neutralizing the risks of visual ambiguity [8], [6].

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The resolution of these machine learning challenges is foundational to securing the integrity of the global herbal medicine supply chain. By bridging the gap between algorithmic innovation and practical field constraints, these technologies ensure that botanical authentication remains a robust safeguard for public health and pharmaceutical safety.

REFERENCES

[1] S. Sendhilkumar, G. S. Mahalakshmi, A. Swaminathan, and K. Sai Anirudh, "Automated Indigenous Plant Recognition and Medicinal Value Extraction System," in Computational Intelligence in Data Science, vol. 717, M. L. Owoc, F. E. Varghese Sicily, K. Rajaram, and P. Balasundaram, Eds., in IFIP Advances in Information and Communication Technology, vol. 717. , Cham: Springer Nature Switzerland, 2024, pp. 368–380. doi: 10.1007/978-3-031-69982-5_28.

[2] S. J. Gamage K.G., A. P.N.P., N. Gamitha Manawadu, and H. De Silva, "AI-Enhanced Medicinal Plant Identification System with Multilingual Social Media Integration," Int. J. Innov. Sci. Res. Technol. IJISRT, pp. 2603–2612, Jun. 2024, doi: 10.38124/ijisrt/IJISRT24MAY1928.

[3] J. Zou and R. Liao, "AI-powered recognition of Chinese medicinal herbs with semantic structure modeling and gradient-guided enhancement," Front. Plant Sci., vol. 16, p. 1672394, Nov. 2025, doi: 10.3389/fpls.2025.1672394.

[4] J. Wu et al., "Artificial intelligence in Traditional Chinese Medicine herbs: a survey," Artif. Intell. Rev., vol. 59, no. 4, p. 120, Feb. 2026, doi: 10.1007/s10462-026-11513-w.

[5] F. Zhao, J. Zhang, Q. Liu, C. Liang, S. Zhang, and M. Li, "Fast Quality Detection of Astragalus Slices Using FA-SD-YOLO," Agriculture, vol. 14, no. 12, p. 2194, Nov. 2024, doi: 10.3390/agriculture14122194.

[6] Y. Liu, H. Liu, L. Li, and Y. Ding, "From image to insight deep learning solutions for accurate identification and object detection of Acorus species slices," Sci. Rep., vol. 15, no. 1, p. 15734, May 2025, doi: 10.1038/s41598-025-00038-x.

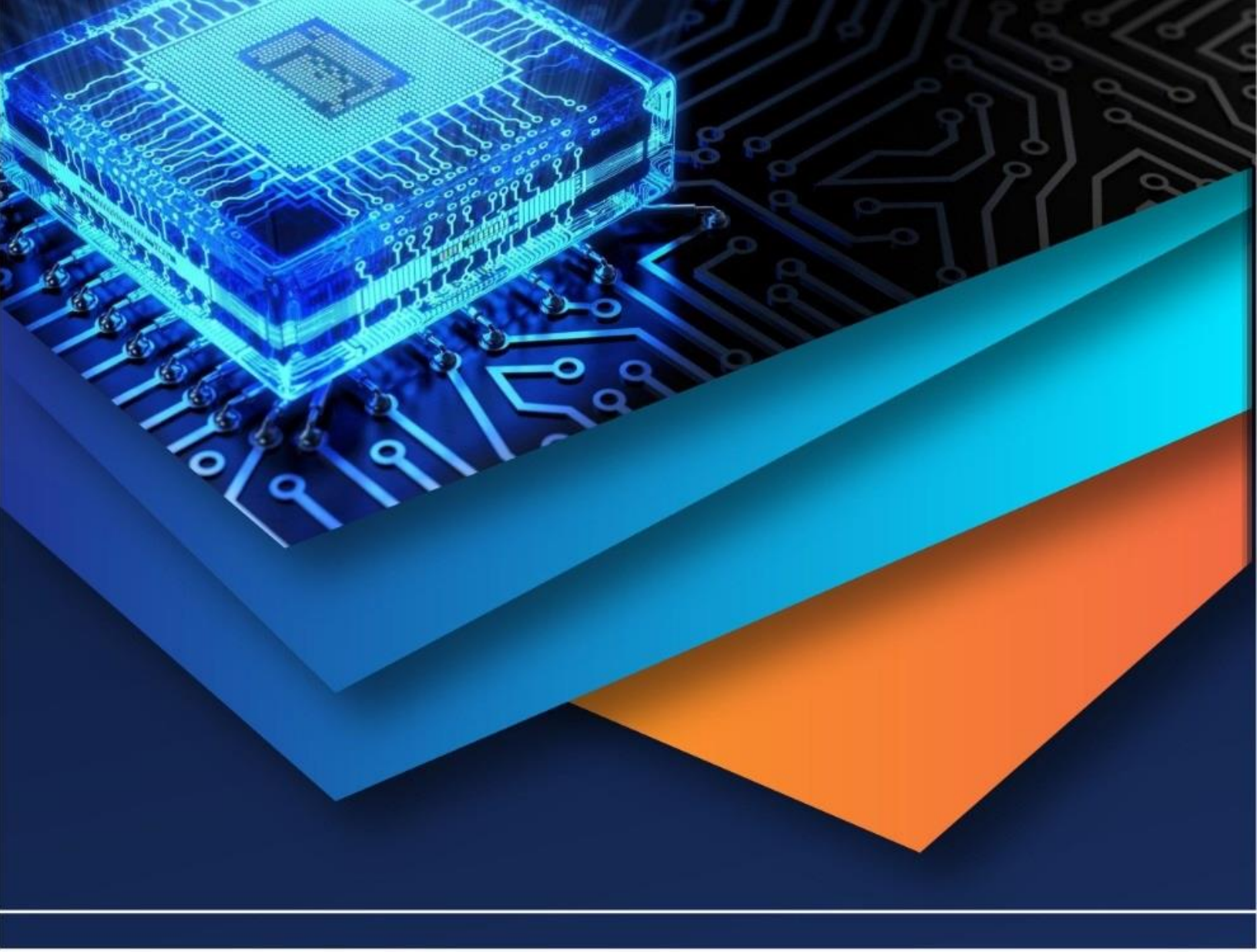
[7] J. Miao, Y. Huang, Z. Wang, Z. Wu, and J. Lv, "Image recognition of traditional Chinese medicine based on deep learning," Front. Bioeng. Biotechnol., vol. 11, p. 1199803, Jul. 2023, doi: 10.3389/fbioe.2023.1199803.

[8] K. Yang, G. Ding, J. Liu, and G. Du, "MGCA-Net: Multi-Granularity Cross-Attention Network with Consistency Learning for Medical Plant Recognition". Jan. 2026, 10.2139/ssrn.6557881.

[9] X. Guan, L. Shi, W. Yang, H. Ge, X. Wei, and Y. Ding, "Multi-Feature Fusion Recognition and Localization Method for Unmanned Harvesting of Aquatic Vegetables," Agriculture, vol. 14, no. 7, p. 971, Jun. 2024, doi: 10.3390/agriculture14070971.

[10] G. Zhu, J. Joo, S. Park, and S. C. Kim, "Toxic Chinese herbal medicine recognition in real-world images via multi-scale and attention-enhanced EfficientNetV2," PLOS One, vol. 21, no. 3, p. e0344262, Mar. 2026, doi: 10.1371/journal.pone.0344262.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)