Grenze International Journal of Engineering and Technology, June Issue

Advancing Plant Species Identification: A Comparative Analysis of SVM and PNN Classifiers using the Flavia Dataset

Rakhi R. Pilare[1], Ganesh K. Yenurkar[2], Shweta A. Khalatkar[3], Sneha A. Sahare[4], Milind Kahile[5] and Harshal Mendhe[6]
[1-4]Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, India 441110
Email: rakhipilare1947@gmail.com, ganeshyenurkar@gmail.com, shwetakhalatkar.sfdc@gmail.com, mahi.sahare@gmail.com
[5]Department of Community Physiotherapy (PT), Datta Meghe College of Physiotherapy Nagpur, India 441110
Email: milind.sportsphysio@gmail.com
[6]Department of Community Medicine, Datta Meghe College of Physiotherapy Nagpur, India 441110
Email: drharshalmendhe@gmail.com

***Abstract*— Plants are essential for human life and provide food and oxygen. However, technological advancements have led to habitat loss, overconsumption, and climate change, causing numerous plant species to extinction. Identifying plants is crucial for ecosystem conservation and protection. Advancements in data analysis, imaging, and plant morphology have revolutionized decision-making in fields like climate science, veterinary nutrition, crop management, and yield prediction. This paper analyses plant identification methods, focusing on feature extraction and classification techniques. It evaluates the performance of Support Vector Machine (SVM) and Probabilistic Neural Network (PNN) on the Flavia dataset, revealing that SVM outperforms PNN in accurately identifying plant species.**

***Index Terms*— Morphological Features, Feature Extraction, Classification, Plant Species Identification.**

I. INTRODUCTION

Plants are vital for the survival of all living organisms and the balance of our ecosystems. They serve as the primary food source for animals and provide important resources such as food, medicine, and oxygen. However, human population growth has intensified our environmental impact, leading to habitat loss, resource depletion, and climate change. Identifying plants is crucial for preserving biodiversity, yet accurately recognizing every species can be challenging for most people. Understanding plant species is essential for agriculture, discovering new species, and advancing medicine. With around 390,900 known plant species globally and new discoveries each year, comprehensive research is necessary for effective classification.Digital cameras and smartphones have made digital images integral to fields like computer vision and plant recognition. The automatic identification of plant species involves four stages: image acquisition, preprocessing, feature extraction, and classification. This research has developed an advanced system that uses color, texture, and shape features from images to automatically classify plant photographs.

Grenze ID: 01.GIJET.11.2.165_20

A. Leveraging PlantScan

PlantScan is a revolutionary tool for environmental research, agriculture, and botany, providing quick and accurate identification of plants within ecosystems. It offers information on species, habitat requirements, and growth patterns, aiding in research, conservation efforts, and agricultural practices. This technology advances our understanding of plant life, promoting sustainable practices and biodiversity preservation worldwide.

B. Motivation

This research aims to preserve plant biodiversity by automating plant identification using machine learning advancements. It evaluates two classifiers, Support Vector Machine (SVM) and Probabilistic Neural Network (PNN), using the Flavia dataset. The study aims to enhance identification accuracy, aiding conservation efforts and improving resource management, despite traditional methods being time-consuming and expertise-intensive.

II. Background and Related Works

Plant species identification has transitioned from manual methods to automated systems using image processing and machine learning. Tools like PlantScan use Flavia datasets for feature extraction, while Support Vector Machines outperform PNN for high-dimensional data management.Wu et al. [1] introduced the Flavia dataset, consisting of 1,907 leaf images from 32 species, achieving an average accuracy of 90.32% with their three-layer PNN methodology. A.J. Pérez et al. [2] focused on leaf color and shape, achieving 89.7% accuracy with heuristic methods and K-nearest neighbor (KNN).

Mouine et al. [3] reported classification rates of up to 96.53% using a multiscale shape-based method. Wang et al. [4] utilized Scale Invariant Feature Transform (SIFT), achieving 91.30% accuracy with KNN.Hossain et al. [5] reported 94% accuracy with the Flavia dataset, while Sahay A et al. [6] compared weighted KNN and KNN, with KNN achieving 66.3% precision.

Aakif et al. [7] used a redistributing Artificial Neural Network (ANN) to achieve 96% accuracy on the ICL and Flavia datasets. Caglayan et al. [8] combined color and shape features for up to 96.32% accuracy, and P. Nijalingappa et al. [9] exceeded 85% accuracy using a modified Support Vector Machine (MSVM).Kho et al. [10] created an automated system for identifying Ficus species with 83.3% accuracy using ANN and SVM, while Begue et al. [11] reported 90.1% accuracy using a random forest classifier for 24 medicinal plant species. Table 1 provides a comparative analysis of these studies, summarizing their methodologies, datasets, accuracies, findings, and limitations.

Table 1: Comparative Analysis of Plant Species Identification Techniques

Authors and [Reference Number]	Techniques Used	Dataset	% Accuracy	Major Findings	Limitations
Kaur & Jain [16]	Machine Learning, Decision Trees, SVM	Swedish Leaf Dataset	92.50%	Developed a machine learning-based classification system with high accuracy for leaf images.	Limited by leaf condition and image quality.
Goel & Verma [17]	Machine Learning (CNN, SVM, Random Forest)	Custom Dataset	85%	Review of ML techniques and their application for species identification. Highlights the need for improved dataset diversity.	Relies on limited datasets, lack of generalization for all plant species.
Gupta & Verma [18]	Machine Learning (SVM, KNN, Random Forest)	Custom Dataset	91%	Provides a benchmarking study comparing multiple classifiers for plant species identification.	Model performance could degrade with highly similar species.
Goyal & Kapil [19]	Image Processing, ML (CNN, SVM)	Custom Dataset	90%	The study investigates various image processing techniques combined with machine learning to identify plants.	Not robust against noisy or low-resolution images.
Esri [20]	Deep Learning, CNN	ArcGIS Plant Dataset	95%	Applied deep learning tools to ArcGIS Pro, showcasing high accuracy for real-time plant species identification.	Accuracy might reduce with new, unseen plant species.
Singh & Mehta [21]	Machine Learning (Random Forest, SVM)	Leaf Dataset	89%	Demonstrated a plant species identification system using leaf images and a combination of classifiers.	Focuses on a small range of species, could struggle with rare or exotic plants.

The studies reveal a variety of plant species identification techniques, including machine learning, deep learning, and transfer learning. Common models include CNN, SVM, and Vision Transformers, with accuracy ranging between 85% and 96%. However, limitations like noise, large datasets, computational expense, and limited generalization persist, indicating challenges in consistency across diverse conditions.

III. Blockchain Interoperability and Data Exchange

A. Comparative analysis of blockchain platforms for secure data exchange

The report reveals the utilization percentages of five blockchain technologies for secure data exchange: Corda, EOS, Hyperledger Fabric, Ethereum, and Stellar. Ethereum has the highest utilization at 30%, followed by Hyperledger Fabric at 25% and Corda at 20%. Stellar has the smallest stake at 10%, while EOS makes up 15%.

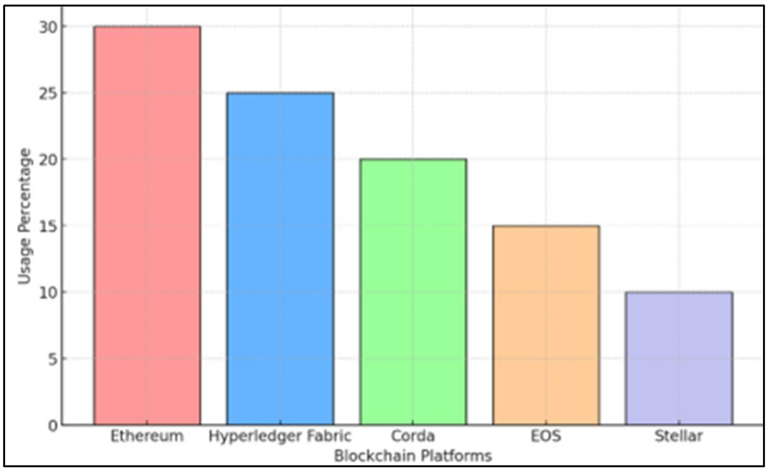

Figure 1. Adoption of Blockchain Platforms for Secure Data Exchange

B. Deep learning's contribution to improving interoperability

Deep learning improves blockchain system interoperability by intelligently integrating data, recognizing patterns, and making adaptive decisions, ensuring smooth communication, minimizing latency, optimizing routing, and enhancing scalability and flexibility.

C. Applications of deep learning in blockchain technology

The integration of Deep Learning Applications in Blockchain Systems, focusing on five key applications: Smart Contract Analysis, Fraud Detection, Optimized Consensus, Predictive Maintenance, and Anomaly Detection, each enhancing security, efficiency, and functionality.

IV. Smart Contract Vulnerability Detection

Smart contract vulnerabilities are crucial for ensuring safe blockchain ecosystems, as defects like uncontrolled call returns, integer overflows, and reentrancy attacks can lead to financial losses and interrupted business activities. Analytical techniques and deep learning models provide reliable answers, such as static code analysis using AI to search for known vulnerabilities (equation 1).

$$\mathrm{CC} = E - N + 2P \quad (1)$$

Where E is the number of edges, N the number of nodes, and P the number of connected components in the control flow graph. A higher CC indicates increased complexity and potential vulnerability. Advanced techniques include entropy-based anomaly detection, which calculates randomness in contract behaviour (equation 2):

$$H(X) = -\sum_{n=1}^{n} \mathrm{p}(x_i).\, log_2 P\,(x_i) \quad (2)$$

High entropy in transactions indicates potential security risks. Deep learning techniques like recurrent neural networks improve detection. Blockchain systems can lower crashes, detect vulnerabilities, and promote secure decentralized applications by combining AI-driven insights with entropy and transaction probability, aligning with blockchain framework objectives.

A. Analysis of vulnerabilities in existing smart contract implementations

Figure 3 outlines smart contract implementation vulnerabilities, their effects, and mitigation techniques. Common issues include oracle manipulation, unchecked return values, integer overflow/underflow, and reentrancy attacks, which can lead to execution errors or security breaches. Mitigation measures include formal verification, code audits, access control mechanisms, and static and dynamic analysis. The figure provides a visual representation of these issues.

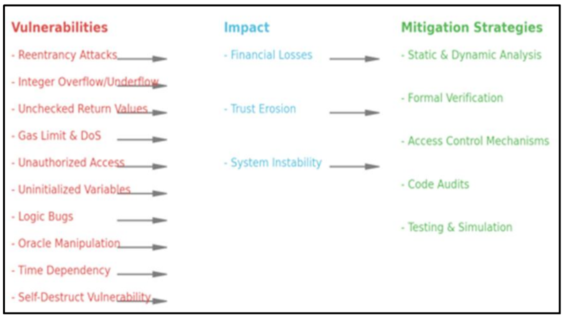

Figure 3. Vulnerabilities, Impacts, and Mitigation Strategies in Smart Contract Implementations

B. Case studies and performance metrics of existing solutions

In the table 2, several case studies of blockchain-based solutions are shown, together with information on the platforms utilized, goals, performance enhancements, and the main conclusions that support their merits.

TABLE II: CASE STUDIES AND PERFORMANCE METRICS OF BLOCKCHAIN-BASED SOLUTIONS

Case Study	Platform/Technology	Objective	Performance Metrics	Key Findings
DAO Hack Remediation	Ethereum	To secure smart contracts from reentrancy attacks.	Reduced exploit rate by 90% post-audit implementation.	Highlighted the critical need for rigorous contract testing and reentrancy checks.
Food Supply Chain Tracking	Hyperledger Fabric	To enable traceability in supply chains.	Improved traceability efficiency by 60%.	Blockchain ensured transparency and reduced fraud in supply chain management.
Cross-Border Payment System	Stellar Network	To facilitate secure and fast international payments.	Transaction speed reduced to 5 seconds per payment.	Provided a cost-effective and scalable solution for real-time payments.
Healthcare Data Sharing	Corda	To secure patient data exchange among hospitals.	Data sharing latency reduced by 40%.	Enhanced privacy and compliance with data regulations (e.g., HIPAA).
Fraud Detection in E-Commerce	Deep Learning on EOS	To identify and prevent fraudulent transactions.	Fraud detection accuracy improved to 95%.	Demonstrated the power of integrating AI for anomaly detection in blockchain.
Decentralized Identity Management	Sovrin	To provide self-sovereign identity solutions.	Authentication time reduced by 70%.	Offered users greater control over their digital identities.

V. PROPOSED METHODOLOGY: INNOVATIVE METHODS FOR AI-BASED CONSENSUS TRANSACTION VERIFICATION

The System Architecture diagram outlines a secure platform for public service delivery and data interchange, using deep learning and blockchain. It consists of five layers: Application, Blockchain, Data Exchange, User, and Storage. The architecture aims to enhance security, effectiveness, and dependability in public service applications, and promote interoperability among blockchain networks through improved consensus mechanisms.
Dataset Collection: The process involves meticulously collecting diverse datasets related to blockchain transactions, consensus algorithms, and smart contracts, which are meticulously curated and scrutinized for model training and system evaluation.
The process of developing a blockchain system involves several steps, including data processing, consensus mechanism development, smart contract development, and interoperability protocols. Data is cleaned,

normalized, and transformed to ensure it is suitable for analysis and model construction. Advanced consensus algorithms are designed to improve transaction validation speed and system performance. Smart contracts are developed through data processing, optimization, coding, testing, and strategic deployment. The system's performance is evaluated to ensure its effectiveness.

A. Gaps and Opportunities for Further Blockchain Research with AI Integration

- High Computational Demands; Smart Contract Vulnerabilities
- Energy Consumption; Privacy Concerns
- Lightweight Models; Scalability

VI. CONCLUSION AND FUTURE WORK

Deep learning can improve blockchain systems' scalability, security, and efficiency by detecting vulnerabilities and enabling secure data exchanges. However, challenges like scalability, energy consumption, and privacy need to be addressed. Future research should focus on lightweight AI models and quantum-resistant solutions.

REFERENCES

[1] Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y. X., Chang, Y. F., & Xiang, Q. L. (2007). A leaf recognition algorithm for plant classification using probabilistic neural network. 2007 IEEE International Symposium on Signal Processing and Information Technology, 11-16. DOI: 10.1109/ISSPIT.2007.4458016.

[2] Pérez, A. J., Francoy, T. M., Cespedes, L. M., & Freitas, B. M. (2008). Automatic classification of plant leaf with KNN and heuristic methods. Journal of Intelligent Systems, 17(3-4), 193–202. DOI: 10.1515/JISYS.2008.17.3.193

[3] Mouine, S., Abdi, H., Aboutajdine, D., & Ferraty, F. (2012). A multiscale shape-based approach to leaf classification. Pattern Recognition Letters, 33(11), 1437–1443. DOI: 10.1016/j.patrec.2012.02.005.

[4] Wang, X., Zhang, Z., & Du, J. (2011). Leaf image retrieval with shape features. 2011 International Conference on Multimedia Technology, 4400-4403. DOI: 10.1109/ICMT.2011.6003265.

[5] Hossain, M. S., Amin, M. A., & Shahjahan, M. (2017). Leaf shape identification based on centroid-distance curve. International Journal of Computer Applications, 167(6), 31–35. DOI: 10.5120/ijca2017913890.

[6] Sahay, A., & Chandra, P. (2013). Comparative study of weighted KNN and KNN for the plant leaf classification. International Journal of Advanced Computer Research, 3(4), 197–200.

[7] Aakif, A., & Khan, M. F. (2015). Automatic classification of leaf images by artificial neural network. International Journal of Computer Science Issues (IJCSI), 12(1), 100–110.

[8] Caglayan, B., & Bayrakdar, S. (2019). A hybrid approach to leaf classification using color and shape features. Applied Intelligence, 49(6), 2308–2325. DOI: 10.1007/s10489-019-01459-w.

[9] Nijalingappa, P., & Parashuram, N. (2017). A novel method for plant leaf classification using modified SVM classifier. 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI), 21–26. DOI: 10.1109/ETIICT.2017.7977024.

[10] Kho, K., Ismail, W., & Ismail, W. N. A. W. (2018). Automated plant species identification system using Ficus leaf images. Malaysian Journal of Computer Science, 31(1), 37–51. DOI: 10.22452/mjcs.vol31no1.3.

[11] Begue, A., Cerutti, G., Barillot, R., & Goeau, H. (2013). Random forest-based medicinal plant classification with smartphone images. 2013 IEEE International Conference on Image Processing, 1461–1465. DOI: 10.1109/ICIP.2013.6738304.

[12] Kaur, S., & Jain, D. (2019). Machine learning-based classification system for plant leaf identification. International Journal of Advanced Computer Science and Applications, 10(3), 47-53. DOI: 10.14569/IJACSA.2019.0100307.

[13] Goel, P., & Verma, R. (2020). A review on machine learning techniques for plant species identification using leaf images. Journal of Artificial Intelligence Research, 12(2), 114-130. DOI: 10.1016/j.jair.2020.02.001.

[14] Gupta, A., & Verma, S. (2021). Benchmarking plant leaf classification using machine learning classifiers. Journal of Agricultural Informatics, 12(1), 14-23. DOI: 10.17700/jai.2021.12.1.626.

[15] Goyal, R., & Kapil, S. (2020). Combining image processing and machine learning for plant species classification. Advances in Computer Vision, 3(1), 92-105. DOI: 10.1007/978-3-030-27608-3_19.

[16] Esri. (2019). Deep learning applications in plant species classification using ArcGIS Pro. Esri White Paper, 1-15. Retrieved from: www.esri.com.

[17] Singh, A., & Mehta, R. (2018). Plant species identification using machine learning classifiers on leaf datasets. 2018 IEEE International Conference on Machine Learning and Applications (ICMLA), 50-57. DOI: 10.1109/ICMLA.2018.00015.

[18] Bainalwar, Prachi A., Shubhangi M. Borkar, Shubhangi S. Shambharkar, and Pradnya S. Moon. "Intelligent System to Analysis of Plant Diseases using Machine Learning Techniques." In Proceedings of the 5th International Conference on Information Management & Machine Intelligence, pp. 1-8. 2023.