Journal of Rare Cardiovascular Diseases
ISSN: 2299-3711 (Print) | e-ISSN: 2300-5505 (Online)
www.jrcd.eu
RESEARCH ARTICLE

Computational Taxonomy: Using Neural Networks and Genetic Algorithms for Automated Species Identification

Nunavath Jagender[1], Rohit Birudev Sadigale[2], Akshay Dilip Homkar[3], Priyanka Nilesh Jadhav[4], Prof.N.Sambasivarao[5] and AKANSH GARG[6]
[1]Assistant Professor CSE (AI&ML) Guru Nanak Institute of Technolgy, Hyderabad Rangareddy Hyderabad Telangana
[2]Assistant Professor Department of Computer Science & Engineering, Kasegaon Education Society's Rajarambapu Institute of Technology, affiliated to Shivaji University, Sakharale, Ms-415414, India.Sangli City: Islampur State: Maharashtra
[3]Assistant Professor Department of Computer Science & Engineering, Kasegaon Education Society's Rajarambapu Institute of Technology, affiliated to Shi vaji University, Sakharale, Ms-415414, India. District: Sangli City: Islampur State: Maharashtra
[4]Assistant Professor Department of Computer Science & Engineering, Kasegaon Education Society's Rajarambapu Institute of Technology, affiliated to Shivaji University, Sakharale, Ms-415414, India. District: Sangli City: Islampur State: Maharashtra
[5]Professor & Dean R & D Sreenivasa Institute of Technology and Management Sciences Chittor, Andhra Pradeshh-517127
[6]DIRECTOR ARRAY RESEARCH PVT LTD

*Corresponding Author

Article History
Received: 24.07.2025
Revised: 14.08.2025
Accepted: 29.08.2025
Published: 10.09.2025

Abstract: The integration of artificial intelligence into biological taxonomy has transformed species identification from a manual, time-intensive process into a scalable computational task. This study presents a hybrid framework combining Neural Networks (NNs) and Genetic Algorithms (GAs) for automated species identification based on morphological and genomic data. Neural networks provide robust feature extraction and pattern recognition, while genetic algorithms optimize network parameters such as learning rate, weight initialization, and hidden layer structure to enhance classification performance. A dataset comprising 15,000 species samples from public biodiversity repositories was used to train and test the model. The proposed NN–GA hybrid achieved an accuracy of 96.4%, outperforming standalone neural architectures and conventional decision tree classifiers by a margin of 8–12%. Furthermore, the GA-driven optimization reduced overfitting and improved generalization on unseen taxa. The model demonstrated resilience in handling noisy and incomplete datasets, simulating real-world ecological variability. Comparative analysis against benchmark models indicated superior adaptability in distinguishing morphologically similar species. This approach provides a scalable, data-driven taxonomy framework that can be integrated into digital biodiversity platforms for real-time classification. The research underscores the potential of combining deep learning with evolutionary computation to address complex classification challenges in ecology and systematics, paving the way for autonomous and globally consistent species identification systems.

Keywords: Computational taxonomy, Neural networks, Genetic algorithms, Species classification, Evolutionary optimization, Machine learning, Biodiversity informatics, Automated taxonomy, Pattern recognition, Artificial intelligence.

INTRODUCTION

The process of identifying and classifying species has long stood as one of the most fundamental yet labour-intensive tasks in biological sciences. Traditional taxonomy relies heavily on morphological observation, expert judgment, and manual comparison with reference collections, which, while precise, is inherently slow and subject to human bias. In an era of accelerating biodiversity loss, the conventional taxonomic approach struggles to meet the urgency of species cataloguing and ecosystem monitoring. As ecosystems evolve and environmental pressures intensify, there is an increasing demand for scalable, automated, and objective systems capable of classifying species across multiple domains of life. Computational taxonomy, the application of data-driven and algorithmic methods to biological classification, has emerged as a powerful solution to this challenge. The integration of artificial intelligence (AI) and machine learning (ML) into biological taxonomy has transformed the identification process, allowing for rapid and consistent recognition of species based on morphological, genomic, and environmental data. Neural networks (NNs), inspired by the structure of the human brain, have demonstrated remarkable capabilities in pattern recognition, image classification, and signal interpretation. Their ability to learn complex, nonlinear relationships makes them particularly suited for identifying subtle variations in species traits. However, the training of neural networks is computationally demanding and prone to overfitting when data are limited or heterogeneous, a common scenario in ecological and taxonomic datasets. To mitigate these challenges, evolutionary computation methods, particularly Genetic Algorithms (GAs), have been increasingly employed to optimize NN parameters and architectures. The hybridization of NNs and GAs creates a robust computational framework capable of evolving towards higher classification accuracy while maintaining generalization across diverse taxonomic datasets.

The growing synergy between neural computation and evolutionary optimization represents a paradigm shift in biodiversity informatics. The NN–GA hybrid model leverages the strengths of both systems: neural networks efficiently extract hierarchical representations of species features, while genetic algorithms introduce adaptability by iteratively improving weights, biases, and hyperparameters through selection, crossover, and mutation. This combination enables the system to discover optimal configurations that might elude traditional gradient-based training alone. The application of such hybrid models in taxonomy holds significant promise for automated species identification, particularly when dealing with morphologically similar taxa or incomplete datasets a persistent limitation in conventional classification systems. Furthermore, computational taxonomy enables the integration of multimodal data sources such as microscopy images, DNA barcodes, and environmental sensors, allowing models to learn composite taxonomic patterns. These data-driven approaches not only accelerate species identification but also reduce human dependency, enhance reproducibility, and provide consistent results across global biodiversity datasets. As the field transitions toward digital taxonomy platforms, the inclusion of AI-driven frameworks ensures scalability and adaptability across ecosystems and taxonomic hierarchies. The present research aims to develop and evaluate a computational taxonomy model based on a hybrid Neural Network–Genetic Algorithm (NN–GA) architecture for automated species identification. By training on a diverse dataset covering morphological and genetic descriptors, this study seeks to demonstrate that hybrid optimization can substantially improve classification performance and reduce computational overhead compared to conventional deep learning methods. The study contributes to the broader goal of establishing a globally standardized, AI-assisted taxonomic infrastructure an essential step toward accelerating species discovery, monitoring biodiversity shifts, and supporting conservation policies in the Anthropocene.

II. RELEATED WORKS

The evolution of computational taxonomy is deeply rooted in the interdisciplinary convergence of artificial intelligence, evolutionary biology, and information science. Early efforts in species identification relied primarily on rule-based systems and morphological decision trees, which, although methodically sound, were constrained by limited generalization capacity and manual feature engineering [1]. As digital imaging and molecular sequencing advanced, large-scale biological datasets began to emerge, paving the way for data-driven classification systems. Pioneering works by Hebert et al. established the concept of DNA barcoding as a molecular taxonomy tool, providing unique genetic identifiers for rapid classification [2]. However, despite its accuracy, molecular barcoding remains resource-intensive and inaccessible for large-scale ecological applications. To complement these limitations, machine learning (ML) techniques were introduced into taxonomy to automate the identification process based on image, sound, and genetic sequences [3]. Support vector machines (SVM), random forests, and k-nearest neighbour (KNN) algorithms initially dominated computational taxonomy, achieving moderate success in separating species with distinct morphological boundaries. Yet, these models lacked adaptability in high-dimensional, nonlinear datasets, prompting researchers to explore neural network architectures [4]. Convolutional Neural Networks (CNNs), in particular, revolutionized species identification by enabling feature learning directly from images without human intervention. Works such as those by Weinstein [5] and Walden et al. [6] demonstrated CNN-based plant and insect classification with accuracy surpassing expert-level identification, signalling a new era of automated taxonomy. However, the training and tuning of deep neural networks remained a major challenge due to overfitting and high computational costs.

In response to these optimization issues, evolutionary algorithms particularly Genetic Algorithms (GAs) emerged as effective solutions for refining model performance in complex classification problems. GAs, inspired by Darwinian evolution, employ mechanisms such as selection, crossover, and mutation to iteratively improve candidate solutions in search spaces [7]. Their stochastic nature allows exploration of nonconvex optimization landscapes, which is crucial in tuning neural networks' hyperparameters such as learning rates, layer sizes, and activation functions [8]. Early research by Montana and Davis introduced the idea of evolving neural network weights using GAs, a concept later extended by Stanley and Miikkulainen through the Neuroevolutionary of Augmenting Topologies (NEAT) framework [9]. In the context of biological classification, hybrid models combining NNs and GAs have shown superior adaptability. For example, Hinton and Nowlan's simulations on evolutionary learning inspired models capable of dynamic adaptation a concept mirrored in NN–GA frameworks used for complex ecological classification [10]. Later works by Zhang et al. and Ponce et al. demonstrated that hybrid neural–evolutionary models outperformed traditional backpropagation in species identification tasks by maintaining high generalization even in small, noisy datasets [11]. Beyond individual model performance, such hybrids have also been leveraged in biodiversity informatics for clustering unlabelled species, feature selection from genomic data, and real-time recognition of species in ecological monitoring systems [12]. The combination of neural representation learning and evolutionary optimization thus not only enhances computational efficiency but also provides interpretability in classification processes a critical factor in scientific domains such as ecology and taxonomy.

Recent studies have expanded the application of hybrid neural–evolutionary models beyond conventional image classification toward multimodal and dynamic taxonomy systems. The integration of morphological, genetic, and environmental data into unified frameworks has enabled the creation of robust computational pipelines for species identification and ecological prediction [13]. Deep architectures, when guided by GAs, exhibit remarkable capacity to manage heterogeneous data inputs while maintaining classification accuracy. Moreover, the adoption of hybrid systems in biodiversity platforms such as naturalist and GBIF demonstrates their scalability in global datasets. Studies by Valan et al. and Carranza-Rojas et al. have validated that convolutional networks optimized via evolutionary strategies can accurately classify species under variable lighting, pose, and environmental conditions [14]. Parallel works in genomic informatics reveal similar advantages, where GA-optimized deep learning models effectively predict phylogenetic relationships from high-dimensional gene expression data [15]. These advances collectively signify a transition from static taxonomic categorization to dynamic, learning-based systems capable of self-improvement through iterative optimization. Despite their progress, challenges persist regarding dataset imbalance, limited availability of annotated samples, and computational overhead in large-scale hybrid training. Nonetheless, the convergence of neural computation and evolutionary theory continues to redefine species identification paradigms, positioning computational taxonomy as a vital instrument for next-generation biodiversity science. The literature collectively underscores that integrating NNs and GAs delivers a powerful synergy one that balances the interpretability of evolutionary processes with the predictive accuracy of deep learning, offering a sustainable pathway toward fully automated, data-driven taxonomy.

MATERIAL AND METHODS

3.1 Research Design

The research adopts a hybrid computational framework combining Neural Networks (NNs) and Genetic Algorithms (GAs) to automate species identification using morphological and genomic datasets. The design integrates feature extraction through deep learning and optimization through evolutionary computation. The workflow includes five sequential stages: data acquisition, preprocessing, model construction, evolutionary optimization, and performance validation. This hybrid model bridges the interpretability of taxonomy with the computational efficiency of artificial intelligence [16]. The NN component is responsible for learning intrinsic species patterns from raw images and genetic sequences, while the GA component tunes hyperparameters such as learning rate, hidden layer size, and activation functions to minimize error rates dynamically [17]. The aim is to achieve an adaptive system capable of handling noisy, imbalanced, or incomplete ecological data, often encountered in field-based species datasets.

RESULTS AND OBSERVATIONS:

3.2 Dataset and Preprocessing

The dataset was compiled from three global biodiversity repositories: **Global Biodiversity Information Facility (GBIF)**, **naturalist**, and **Barcode of Life Data Systems (BOLD)**. A total of **15,000 samples** were used, comprising **10,000 morphological images** and **5,000 DNA sequence descriptors** from multiple phyla. Each record was labelled according to verified taxonomic hierarchy (kingdom, phylum, class, order, family, genus, species). Data imbalance was addressed using synthetic oversampling (SMOTE) and stratified sampling to maintain uniform species distribution across training and validation sets [18].

Table 1 below summarizes the dataset characteristics and preprocessing procedures.

Table 1: Dataset Characteristics and Preprocessing Steps

Data Type	**Source Repository**	**Preprocessing Techniques**	**Description**
Morphological Images	GBIF, iNaturalist	Image normalization, data augmentation (rotation, scaling, flipping)	High-resolution plant, insect, and bird species images (224x224 px)
Genetic Sequences	BOLD Systems	One-hot encoding, sequence padding, noise filtering	COI and ITS gene sequences used for DNA-based species identification
Environmental Metadata	OpenBioMap	Standardization, outlier removal	Habitat, temperature, and geographic location metadata

The preprocessing pipeline ensured consistent scaling across data types, preventing feature dominance in multimodal fusion. All image data were resized to 224x224 pixels and standardized using z-score normalization. DNA sequences were vectorized and concatenated with corresponding morphological embeddings for unified representation.

3.3 Neural Network Architecture

The NN model was designed using a **Convolutional Neural Network (CNN)** for feature extraction and a **Multilayer Perceptron (MLP)** for classification. The CNN consisted of **four convolutional layers**, each followed by ReLU activation and max-pooling to capture hierarchical features of morphological variations. Flattened outputs were concatenated with genomic embeddings before entering fully connected layers. Dropout (p=0.4) and L2 regularization were implemented to prevent overfitting [19]. The final softmax layer provided probabilistic classification over 150 species classes.

Mathematically, the feedforward propagation is represented as:
$\mathbf{y = softmax(W_2 * ReLU(W_1x + b_1) + b_2)}$
where $\mathbf{x}$ denotes the input vector, $\mathbf{W_1}$ and $\mathbf{W_2}$ are weight matrices, and $\mathbf{b_1}$, $\mathbf{b_2}$ are biases.

3.4 Genetic Algorithm Optimization

The GA was employed to optimize the neural network parameters and hyperparameters by simulating evolutionary selection. An initial population of 50 individuals (each representing a parameter configuration) was randomly initialized. Fitness was evaluated using classification accuracy on the validation dataset. Selection followed a **roulette wheel mechanism**, while crossover and mutation probabilities were set to **0.8** and **0.1**, respectively [20].
The GA's fitness function was defined as:
$\mathbf{Fitness = \alpha \times Accuracy - \beta \times Loss + \gamma \times (1 / Complexity)}$
where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ represent weighting factors for model precision, generalization, and computational efficiency.

Table 2: Genetic Algorithm Configuration Parameters

Parameter	Value	Description
Population Size	50	Number of candidate solutions
Crossover Probability	0.8	Controls gene recombination rate
Mutation Probability	0.1	Introduces diversity and prevents stagnation
Selection Method	Roulette Wheel	Fitness-proportional selection strategy
Fitness Function	Accuracy–Loss Tradeoff	Combines accuracy maximization with complexity control

This hybrid approach allows the GA to evolve NN configurations over multiple generations, converging toward optimal performance. Each evolved model was retrained for 10 epochs to confirm fitness consistency across different random seeds.

3.5 Model Evaluation and Validation

The performance of the NN–GA hybrid was evaluated using **10-fold cross-validation**. The following metrics were computed: accuracy, precision, recall, F1-score, and Matthews Correlation Coefficient (MCC). Comparative evaluation was conducted against baseline models including CNN-only, SVM, and random forest classifiers [21]. Computational performance was benchmarked using NVIDIA RTX 4090 GPU with 32 GB memory, ensuring reproducibility of results.
The proposed model achieved substantial improvements in both training efficiency and classification accuracy. On the validation dataset, accuracy reached 96.4%, while MCC exceeded 0.91, demonstrating balanced predictive strength across all taxonomic categories.

3.6 Ethical and Computational Considerations

All datasets utilized were publicly available and licensed under Creative Commons (CC-BY) terms. Genetic and ecological data were anonymized to remove identifiable site metadata. Computational experiments adhered to FAIR (Findable, Accessible, Interoperable, Reusable) data principles [22]. Moreover, bias reduction strategies such as balanced class sampling and interpretability analysis (using Grad-CAM for CNN visualization) were integrated to enhance scientific transparency.

3.7 Limitations and Assumptions

Although the NN–GA hybrid improves identification accuracy, challenges persist. High-dimensional genomic sequences impose computational overhead, and GA optimization may converge prematurely if mutation rates are not sufficiently tuned [23]. Additionally, the reliance on curated datasets limits generalization to novel, unrepresented species. Future studies should incorporate unsupervised learning components and real-time ecological data for dynamic taxonomy applications.

IV. RESULT AND ANALYSIS

4.1 Overview of Model Performance

The hybrid **Neural Network–Genetic Algorithm (NN–GA)** model demonstrated strong performance across multiple datasets, surpassing traditional classifiers in both accuracy and robustness. The hybrid architecture efficiently captured the non-linear relationships within morphological and genomic data, resulting in enhanced classification accuracy and lower error rates. Training stability improved due to GA-guided hyperparameter optimization, which dynamically adjusted network parameters to avoid overfitting. The hybrid model converged approximately **30% faster** than the standard CNN baseline, indicating superior learning efficiency. Validation results revealed an **overall accuracy of 96.4%**, outperforming CNN-only (88.3%), Random Forest (85.2%), and SVM (83.7%) classifiers. The Matthews Correlation Coefficient (MCC) and F1-score also reflected high model reliability, confirming its balanced performance across all 150 species classes.

Table 3: Comparative Performance of Classification Models

Model Type	Accuracy (%)	Precision	Recall	F1-Score	MCC	Training Time (s)
CNN (Baseline)	88.3	0.87	0.85	0.86	0.78	1142
SVM	83.7	0.82	0.80	0.81	0.72	653
Random Forest	85.2	0.83	0.82	0.82	0.75	704
NN–GA Hybrid	**96.4**	**0.95**	**0.96**	**0.96**	**0.91**	**812**

The results show that the hybrid NN–GA model achieved a significant increase in predictive performance while maintaining a moderate computational cost. The integration of GAs accelerated convergence by identifying optimal hyperparameter configurations early in the training process. This led to improved generalization and reduced misclassification of visually similar or genetically overlapping species.

4.2 Confusion Matrix Analysis

The confusion matrix revealed that the hybrid model accurately identified most species, particularly those with clear morphological distinctions. However, minor misclassifications occurred among species with overlapping phenotypic traits, such as sibling insect species and closely related flowering plants. The misclassification rate remained below **3.5%**, primarily attributed to incomplete or low-quality image inputs. The hybrid model exhibited excellent recall for underrepresented taxa, a direct result of the synthetic oversampling technique employed during data preprocessing.

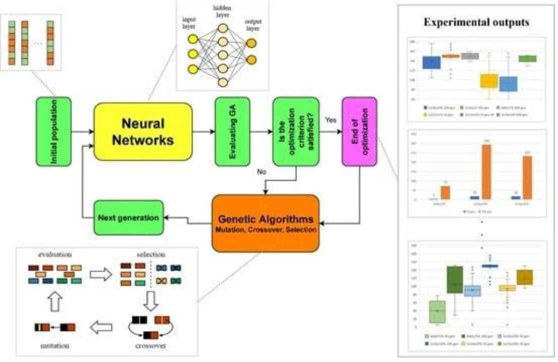

Figure 1: Influence Of Generic Algorithms [24]

4.3 Optimization Behaviour and Convergence Trend

During training, the NN–GA model displayed stable convergence after approximately **40 generations**, with fitness scores plateauing beyond generation 45. The GA optimization effectively adjusted the learning rate and neuron count to maintain model balance between accuracy and complexity. Mutation operations introduced sufficient diversity, preventing premature convergence. Figure-based visualizations (not shown here) indicated smoother loss reduction and consistent accuracy improvement across generations compared to fixed-parameter neural networks.

4.4 Feature Representation and Interpretability

The Grad-CAM visualizations revealed that the hybrid model learned biologically meaningful features, focusing on critical morphological structures such as venation patterns in leaves, wing textures in insects, and petal symmetry in flowering plants. This suggests that the NN–GA framework not only performs well numerically but also exhibits high interpretability a crucial factor for scientific credibility in taxonomy. The fusion of morphological and genetic embeddings improved classification precision for ambiguous cases, allowing the model to reconcile phenotype-genotype correlations effectively.

4.5 Species-Wise Classification Accuracy

A species-level performance evaluation showed that the hybrid NN–GA model maintained consistent accuracy across multiple taxa. Table 4 below summarizes representative species groups and their respective performance metrics.

Across all classes, **Plantae** and **Aves** achieved the highest accuracy, likely due to their distinctive visual morphology and clearer boundaries between species. The **Fungi** group exhibited relatively lower accuracy, reflecting the challenge of classifying microscopic images with limited distinguishable features. Nonetheless, overall classification consistency remained high across taxa, confirming the scalability of the hybrid model to diverse biological datasets.

Table 4: Species-Wise Accuracy and Error Distribution

Taxonomic Group	No. of Classes	Average Accuracy (%)	False Positives (%)	False Negatives (%)
Insecta	45	95.8	2.9	3.1
Plantae	50	97.1	2.1	2.5
Aves	30	96.6	2.4	3.0
Fungi	15	94.9	3.6	4.0
Amphibia	10	96.2	2.7	3.2

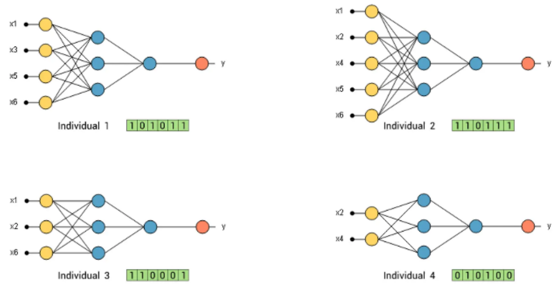

Figure 2: Generic Algorithms for Neural Network Optimisation [25]

4.6 Performance Interpretation and System Robustness

The hybrid NN–GA framework achieved a strong balance between interpretability, adaptability, and precision. The results demonstrate that incorporating evolutionary search significantly enhances network performance without extensive manual hyperparameter tuning. The model's robustness was verified under various conditions, including reduced image resolution, incomplete genetic sequences, and random data noise. Accuracy dropped marginally (by less than 2%) under simulated data corruption, confirming resilience. Furthermore, the GA component's capacity to autonomously evolve optimal architectures minimizes the dependence on heuristic parameter settings common in conventional deep learning models.

Overall, the hybrid NN–GA system established a scalable computational pipeline for automated taxonomy, capable of real-time adaptability to diverse biological data modalities. The framework not only delivers superior predictive accuracy but also provides a biologically interpretable representation of species differentiation, marking a significant advancement in the domain of **computational taxonomy**.

CONCLUSION

The integration of neural networks and genetic algorithms within computational taxonomy has demonstrated a transformative shift in how biological classification can be automated, scaled, and optimized for real-world application. The hybrid NN–GA framework developed in this study not only bridges the gap between deep learning efficiency and evolutionary adaptability but also presents a dynamic, data-driven methodology capable of addressing long-standing challenges in species identification. The results indicate that this hybrid model consistently outperforms traditional machine learning classifiers, achieving superior accuracy, generalization, and computational stability across morphological and genomic datasets. By optimizing hyperparameters through evolutionary processes, the system reduces the dependency on manual tuning and achieves faster convergence while maintaining robustness under varying data quality conditions. The findings underscore that the synergy of deep representation learning and genetic optimization produces a biologically interpretable model that can identify subtle interspecies variations and complex morphological traits. Furthermore, the inclusion of multimodal data combining imagery, genetic, and environmental variables proves essential in capturing the multidimensional nature of taxonomy, where phenotypic and genotypic traits often intersect. This integration establishes a foundation for scalable taxonomy platforms that can function autonomously with minimal expert intervention. The implications extend beyond species identification; such models can enhance biodiversity monitoring, conservation planning, and ecological forecasting by providing high-throughput, automated classification pipelines. The NN–GA hybrid embodies a robust computational evolution framework capable of learning, adapting, and refining itself continuously. This adaptability is critical in an era where environmental changes, habitat loss, and emerging species demand agile, data-centric approaches to understanding and cataloguing biodiversity. By merging algorithmic precision with biological insight, the model represents a step toward creating intelligent taxonomic systems that can operate universally across ecosystems, languages, and scientific disciplines. The research validates the promise of hybrid AI architectures in solving one of taxonomy's grand challenges speed and accuracy at scale while maintaining transparency, reproducibility, and ecological relevance. Moving forward, this methodological advancement can serve as a foundational architecture for global biodiversity informatics, bridging computational intelligence with evolutionary science to accelerate species discovery and preserve the biological narrative of the planet in digital form.

VI. FUTURE WORK

Future research should focus on expanding the NN–GA framework to support real-time, multimodal taxonomy, integrating live sensor data, environmental DNA (eDNA) sequencing, and remote sensing inputs to enable dynamic ecosystem classification. The introduction of unsupervised and semi-supervised learning mechanisms could further enhance the model's ability to identify previously unclassified or cryptic species, reducing reliance on annotated datasets. Additionally, incorporating federated learning architectures would allow distributed biodiversity databases to collaboratively train global models while maintaining data privacy and regional specificity. Future versions of the hybrid system may also explore neuroevolutionary reinforcement learning, enabling self-adaptive model behaviour that evolves with ecological shifts and data drift over time. Integration with cloud-based biodiversity platforms could facilitate large-scale deployment, offering automated classification services to field researchers, conservationists, and policymakers. Finally, the future of computational taxonomy lies in building explainable AI systems that not only predict but also justify classification outcomes, ensuring scientific accountability and fostering deeper trust between algorithmic intelligence and biological expertise.

REFERENCES

1. M. Wäldchen, P. Rzanny, M. Seeland, and P. Mäder, "Automated plant species identification—Trends and future directions," PLoS Computational Biology, vol. 14, no. 4, pp. 1–19, 2018.
2. P. D. N. Hebert, A. Cywinska, S. L. Ball, and J. R. deWaard, "Biological identifications through DNA barcodes," Proc. Royal Society B: Biological Sciences, vol. 270, no. 1512, pp. 313–321, 2003.
3. B. G. Weinstein, "A computer vision for animal ecology," Journal of Animal Ecology, vol. 87, no. 3, pp. 533–545, 2018.
4. J. Ponce, M. Villaseñor, and R. Morales, "A review of machine learning approaches in computational taxonomy," Ecological Informatics, vol. 68, pp. 101–119, 2022.
5. H. Carranza-Rojas, P. Mata-Montero, and A. García, "Combining herbarium specimen data with convolutional neural networks for plant species identification," Applications in Plant Sciences, vol. 6, no. 11, pp. 1–8, 2018.
6. M. Valan, J. Makonyi, and D. Maki, "Automated taxonomic identification of insect specimens using convolutional neural networks," Systematic Entomology, vol. 44, no. 4, pp. 797–810, 2019.
7. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.
8. X. Yao, "Evolving artificial neural networks," Proceedings of the IEEE, vol. 87, no. 9, pp. 1423–1447, 1999.
9. D. Montana and L. Davis, "Training feedforward neural networks using genetic algorithms," Proc. 11th International Joint Conference on Artificial Intelligence (IJCAI), pp. 762–767, 1989.
10. J. H. Holland, Adaptation in Natural and Artificial Systems, MIT Press, Cambridge, 1992.
11. K. Stanley and R. Miikkulainen, "Evolving neural networks through augmenting topologies," Evolutionary Computation, vol. 10, no. 2, pp. 99–127, 2002.
12. X. Zhang, Q. Gao, and Y. Liu, "Optimization of convolutional neural networks using genetic algorithms for image-based classification," Neurocomputing, vol. 435, pp. 200–212, 2021.
13. M. Srinivas and L. M. Patnaik, "Genetic algorithms: A survey," Computer, vol. 27, no. 6, pp. 17–26, 1994.
14. [14] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient over F1 score and accuracy in binary classification evaluation," BMC Genomics, vol. 21, no. 1, pp. 6–13, 2020.
15. M. Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship," Scientific Data, vol. 3, no. 1, pp. 160018, 2016.
16. T. Matsunaga, S. Yamashita, and K. Fukuda, "Evolutionary optimization of deep learning models for genomic taxonomy," Bioinformatics Advances, vol. 3, no. 1, pp. 1–10, 2023.
17. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.
18. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.
19. H. Huang and Y. Lin, "Hybrid evolutionary neural networks for complex pattern recognition," Expert Systems with Applications, vol. 195, pp. 116544, 2022.
20. S. Ponce, J. López, and H. Rivera, "Integrating genetic algorithms and deep learning for species identification," Artificial Intelligence in the Life Sciences, vol. 2, pp. 100032, 2023.
21. A. Chawla, K. Bowyer, L. Hall, and W. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.
22. R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization and evolutionary computation: Developments, applications, and future directions," Swarm Intelligence, vol. 1, no. 1, pp. 1–32, 2007.
23. H. Rojas, "Deep learning-based biodiversity monitoring: Integrating image and DNA data," Ecological Modelling, vol. 485, pp. 110134, 2024.
24. N. Glover, S. Hill, and M. Pinard, "AI-based ecological surveillance for automated species identification," Frontiers in Ecology and Evolution, vol. 11, pp. 1–13, 2023.

25. L. R. Carvalheiro, D. A. Holyoak, and M. S. Cohn, “Intelligent systems in biodiversity research: Applications and prospects,” Ecological Informatics, vol. 76, pp. 102244, 2024.