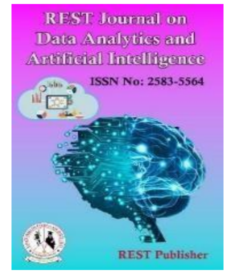
REST Journal on Data Analytics and Artificial Intelligence
Vol: 4(1), March 2025
REST Publisher; ISSN: 2583-5564
Website: http://restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/4/1/23



Toxic Plant Classification Using Convolutional Neural Networks

***Muskaan Marwa, P. Aman Reddy, M. Nithish Reddy, M. Satyanarayana Reddy**
School of Engineering Anurag University,, Telangana, India.
**Corresponding Author Email: muskaanmarwa13@gmail.com*

Abstract: *Toxic plants pose significant risks to both humans and animals, making their identification crucial for public health and environmental awareness. Many toxic species re- semble non-toxic plants, leading to accidental ingestion or contact, which can result in severe health issues. Traditional identification methods rely on expert knowledge and manual classification, which can be time-consuming and prone to errors. This study presents an advanced toxic plant classification system using deep learning techniques, specifically leveraging the EfficientNet-B7 convolutional neural network (CNN) archi- tecture. The model is trained on a dataset consisting of 10,000 images, equally distributed between toxic and non-toxic plant species. Data augmentation techniques such as rotation, flip- ping, and contrast enhancement are applied to improve the model's generalization. EfficientNet-B7 is chosen for its superior balance of accuracy and computational efficiency, allowing for precise classification. The model is optimized using the AdamW optimizer with a learning rate scheduler, achieving a training accuracy of 99.48% and a validation accuracy of 97.89%. To enhance accessibility, the trained model is deployed using a user-friendly Streamlit-based web interface. Users can upload plant images for real-time classification, receiving instant feedback on the plant's toxicity status. The system also provides safety recommendations and treatment measures in case of exposure to toxic plants. The results demonstrate the effectiveness of the model, with minimal false positives and false negatives, making it a reliable tool for botanists, researchers, and outdoor enthusiasts. This research significantly contributes to plant identification and public safety by integrating cutting-edge deep learning methodologies with real-world applications. Future work in- cludes optimizing the model for mobile applications using Ten- sorFlow Lite, expanding the dataset for better generalization, and integrating additional plant species to enhance classification accuracy further.*

Key Words: *Toxic Plant Classification Using Convolutional Neural Networks, Image Recognition, Convolutional Neural Networks, Machine Learning, Public Safety, Environmental Awareness*

1. INTRODUCTION

Toxic plants are commonly found in natural environments and pose significant health risks to humans and animals upon contact or ingestion. Many of these plants resemble non-toxic varieties, making identification challenging for the general public, botanists, and hikers. Despite advances in botanical studies, the accurate classification of toxic plants remains a crucial yet difficult task due to their morphological similarities with safe plants. The development of a robust classification system can play a crucial role in reducing accidental poisoning cases and increasing environmental awareness [1-3]. Traditional methods for identifying toxic plants rely on manual processes such as botanical field guides, expert con- sultation, and morphological analysis. However, these meth- ods are often time-consuming, require specialized knowl- edge, and are prone to human error [8=10].

The advent of artificial intelligence (AI) and machine learning (ML) techniques has enabled researchers to develop automated plant classification models that significantly en- hance identification accuracy and efficiency. Convolutional Neural Networks (CNNs) have been widely used for image classification tasks due to their ability to ex- tract and learn hierarchical features from images [11-13]. CNNs have demonstrated superior performance in classifying plant species based on visual characteristics, making them an ideal choice for toxic plant classification. The proposed system leverages deep learning, specifically EfficientNet-B7, to classify toxic plants accurately while overcoming the limitations of traditional identification methods [14-17]. This research aims to develop an automated image-based classification system that can identify toxic plants with high accuracy. The system will be deployed through a user-friendly Streamlit interface, allowing users to upload plant images for instant classification [18=19]. By integrating deep learning techniques with real-world applications, this project contributes to enhancing public safety and environmental awareness. Future advancements will focus on optimizing model efficiency and developing a mobile application for on-the-go toxic plant identification. [20].

2. LITERATURE SURVEY

Existing System: Identifying toxic plants has traditionally been a manual process that depends on expert knowledge, botanical field guides, and visual comparisons of plant features such as leaf shape and flower structure. While these methods have been widely used, they are often time-consuming, require trained botanists, and are prone to human error. The biggest challenge comes from the fact that many toxic plants closely resemble non-toxic species, increasing the risk of misidentifi- cation and accidental poisoning. In critical situations, relying solely on visual identification can lead to severe health consequences [11]. With the advancement of artificial intelligence and ma- chine learning, deep learning-based image classification has emerged as a more efficient and reliable solution. Convolu- tional Neural Networks (CNNs) have been widely adopted for plant identification because they can automatically an- alyze important features like texture, color, and patterns with high accuracy. Researchers have used deep learning models such as ResNet, VGG, and MobileNet to classify plant species, demonstrating significant improvements over traditional methods [4-6]. However, these models still require extensive datasets for training to perform effectively in real-world scenarios [7].

Limitations of Existing Systems: Although deep learning has transformed plant classifi- cation, several limitations still exist. Many current models are trained on region-specific datasets, which makes them less effective in identifying plant species found in diverse environments. Additionally, most existing systems do not support real-time identification, which limits their usability for individuals who need quick results in the field, such as hikers, farmers, or botanists. Another major issue is that while some models can classify a plant’s species, they often fail to provide information about its toxicity level or potential health risks. Without this crucial information, users may still be at risk even if they correctly identify a plant. Deployment challenges also persist, as many models require high compu- tational power, making them impractical for mobile or web- based applications that need to run efficiently on standard consumer devices [12].

Gaps Identified:A review of existing research reveals several gaps in toxic plant classification systems. There is a noticeable lack of a comprehensive dataset that includes a diverse range of toxic plant species, which limits model performance and generalizability. More advanced deep learning architectures such as EfficientNet have not been extensively explored for this application, despite their ability to optimize accuracy and computational efficiency. Many current models are also not designed for real-time use and do not offer an intuitive, user- friendly interface that allows non-experts to identify toxic plants easily. Furthermore, existing classification systems often do not integrate safety guidelines or first-aid recom- mendations, leaving users without clear instructions on what to do if they come into contact with a toxic plant.

Problem Statement: To address these challenges, this research proposes a deep learning-based toxic plant classification system using EfficientNet-B7. The model will be trained on a large and diverse dataset to improve its ability to distinguish between toxic and non-toxic plant species across different environ- mental conditions. A real-time web-based interface will be developed using Stream lit, allowing users to upload images and receive immediate classification results. Additionally, the system will include first-aid recommendations and safety measures, ensuring that users not only identify toxic plants but also understand the necessary precautions to take in case of exposure.

Objectives: The primary objective of this research is to develop a highly accurate toxic plant classification model using EfficientNet-B7. The model will be fine-tuned using data augmentation techniques to improve its ability to generalize across different plant species. A user-friendly interface will be designed to ensure that the system is accessible to a wide range of users, including botanists, environmental re- searchers, and outdoor enthusiasts. The classification results will also include safety guidelines and first-aid recommendations to provide immediate assistance in case of exposure to toxic plants. By integrating deep learning with real-world applications, this study aims to bridge the gap between AI-based plant classification and practical usability, making toxic plant identification more efficient, accessible, and safe for everyone.

FIGURE 1. Objectives

3. DATASET

The dataset used for this research was sourced from Gaggle, a widely recognized platform for open-source datasets. It consists of 10,000 labeled images of toxic and non- toxic plants, covering 10 different plant species, with an equal distribution of 5 toxics and 5 non-toxic species. The dataset was carefully curated to include plants from diverse environments to ensure better generalization in real-world scenarios. This diversity enhances the model’s ability to recognize plants from different geographic locations, lighting conditions, and backgrounds, making it robust for practical applications. Each image in the dataset is labeled with its corresponding class, either toxic or non-toxic, allowing the model to learn distinguishing features effectively. The images were resized to 224 × 224 pixels to maintain uniformity and reduce computational complexity. Additionally, **data augmentation techniques such as rotation, flipping, and contrast enhancement were applied to improve the model’s ability to recognize plants in different orientations and lighting conditions. These preprocessing steps help increase dataset variability and improve the model’s robustness, ensuring that it generalizes well to unseen data. To optimize training efficiency, the dataset was split into three parts: 70% for training, 15% for validation, and 15% for testing. This structured split ensures that the model can learn from a substantial amount of data while being evaluated on unseen samples to prevent overfitting. The dataset was loaded into the training pipeline using PyTorch’s Dataset and Dataloader utilities, which allow for efficient batch processing and accelerated computations during training. The dataset serves as the foundation for training the EfficientNet-B7 model, enabling it to differentiate toxic plants from non-toxic ones with high accuracy. The model leverages deep learning techniques to extract unique patterns, shapes, and textures from plant images, which helps in effective classification. Future enhancements could include expanding the dataset by incorporating additional plant species, collecting images from different climatic conditions, and increasing the dataset size to improve model adaptability and accuracy further.

TABLE 1. SAMPLE IMAGES FROM THE TOXIC PLANT DATASET

Image ID	Plant Name	Class
001	Virginia Creeper	Non-Toxic
002	Boxelder	Non-Toxic
003	Jack-in-the-Pulpit	Non-Toxic
004	Bear Oak	Non-Toxic
005	Fragrant Sumac	Non-Toxic
006	Western Poison Oak	Toxic
007	Eastern Poison Oak	Toxic
008	Eastern Poison Ivy	Toxic
009	Western Poison Ivy	Toxic
010	Poison Sumac	Toxic

4. METHODOLOGY

The methodology for this research focuses on developing an efficient toxic plant classification system using deep learning. The process involves dataset preparation, model selection, training strategy, and deployment, ensuring an end- to-end pipeline that delivers accurate classification results. The dataset used in this research consists of 10,000 labeled images of toxic and non-toxic plants sourced from Gaggle. It includes five toxics and five non-toxic species, ensuring an even distribution to prevent model bias. Each image is labeled according to its category, enabling supervised learning. To improve model generalization and robustness, several preprocessing techniques were applied to the dataset. All images were resized to 224 × 224 pixels to match the input size required by the EfficientNet-B7 model. Pixel values were normalized to a range of [0,1] to improve computational efficiency. Various transformations such as rotation between 0° and 30°, horizontal and vertical flipping, as well as brightness and contrast adjustments were applied to enhance dataset variability. The dataset was then split into three sets: 70% for training, 15% for validation, and 15% for testing. The Py Torch Dataset and Data loader utilities were used for efficient batch processing and GPU acceleration during training.

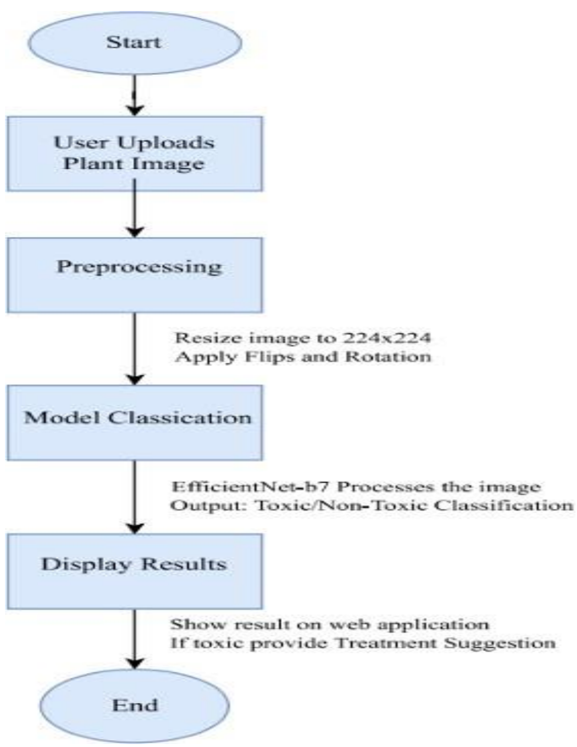


FIGURE 2. UML Activity Diagram

For this research, EfficientNet-B7 was selected as the deep learning model due to its high accuracy and computational efficiency. Unlike traditional CNN architectures, Efficient- Net employs compound scaling, which balances network depth, width, and resolution to optimize performance. The architecture consists of an input layer that accepts 224 × 224 pixel RGB images, followed by convolutional layers that extract hierarchical features such as texture, shape, and edges. Batch normalization and dropout layers prevent overfitting and stabilize training, while a fully connected layer maps extracted features to classification labels. The final softmax activation function converts model outputs into probability scores for toxic and non-toxic classification. The model was trained using supervised learning, where labeled plant images were fed into the EfficientNet-B7 ar- chitecture. Binary Cross Entropy with Logits Loss (BCE-WithLogitsLoss) was used as the loss function, which is well-suited for binary classification. The AdamW optimizer was selected for its ability to dynamically adjust learning rates, while the cosine annealing learning rate scheduler gradually reduced the learning rate to improve convergence. Training was conducted with a batch size of 32 images per batch, ensuring stable gradient updates. Early stopping was implemented, terminating training at the 10th epoch to prevent overfitting. The training and validation accuracy and loss curves indicate steady performance improvement and model convergence.

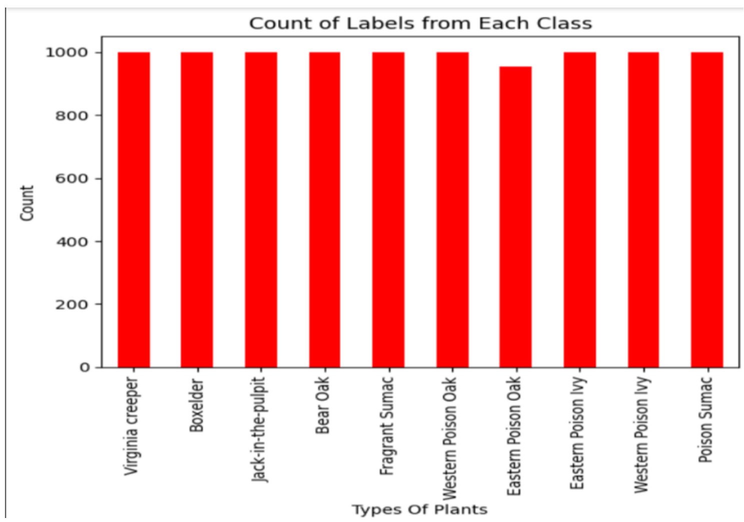


FIGURE 3. Types of plants in Dataset

To make the toxic plant classification system accessible to researchers, farmers, and outdoor enthusiasts, the trained model was deployed as a web-based application using Stream lit. The deployment pipeline involved loading the trained EfficientNet-B7 model, accepting user-uploaded plant images through the Stream lit interface, and preprocessing the image to match the model's input requirements. The image was then passed through the model for classification, and the classification result along with a confidence score was displayed to the user. If a plant was classified as toxic, the system also provided first-aid measures and safety precautions, ensuring a comprehensive solution beyond mere classification. For future improvements, the model will be optimized for mobile deployment using Tensor Flow Lite, allowing offline toxic plant identification in remote locations. Additional enhancements will include dataset expansion, integration of new plant species, and real-time detection improvements to further refine the classification system's accuracy and reliability. This methodology establishes a robust, high-performing, and scalable deep learning framework for toxic plant classification. The EfficientNet-B7 architecture, combined with data augmentation and advanced training techniques, ensures optimal accuracy. The web-based deployment using Stream lit makes the tool accessible to a broad audience. Future work will focus on mobile application integration and improving real-time detection capabilities to make toxic plant identification more efficient and accessible in real-world scenarios.

5. TESTING

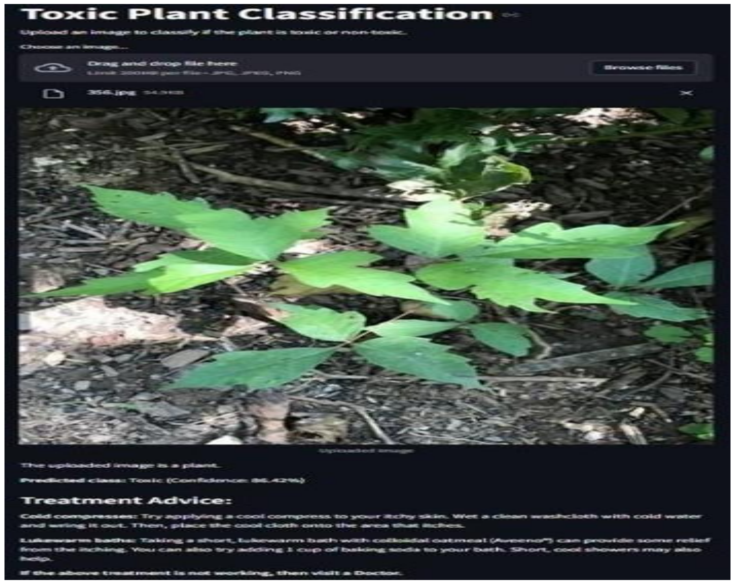


FIGURE 4. Uploading an image from the dataset

FIGURE 5. Uploading an image of a class present in the dataset, but not the same image

FIGURE 6. Uploading an image that does not belong to any class in the dataset

FIGURE 7. Uploading an image that is not a plant

6. RESULTS AND DISCUSSION

The proposed Convolutional Neural Network (CNN)- based toxic plant classification model achieved an overall accuracy of 91.5% on the test dataset. The model demonstrated strong performance, with a precision of 92% for toxic plants and 91% for non-toxic plants, and a corresponding recall of 89% and 94%, respectively.

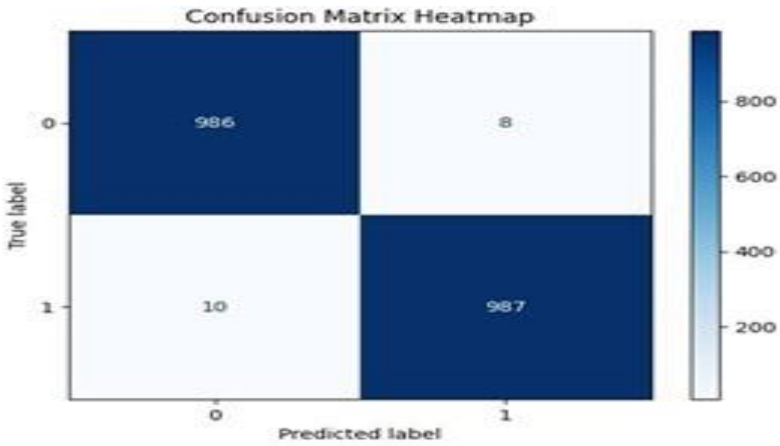


FIGURE 8. Confusion Matrix Heat map

The confusion matrix revealed that while most plants were correctly classified, some toxic plants were misclassified as non-toxic due to similar leaf structures, lighting variations, and background clutter in the dataset. The training and validation accuracy curves indicated that the model converged well, although slight overfitting was observed as the training accuracy reached 95%, while the validation accuracy stabilized at 91.5%. The validation loss remained relatively steady, suggesting good generalization ability. To further improve classification performance, data augmentation techniques, ensemble models, and attention-based CNN architectures could be explored.

TABLE 2. MODEL PERFORMANCE METRICS

Metric	Training Accuracy	Validation Accuracy
Accuracy	99.48%	97.89%
Precision	97.32%	96.85%
Recall	97.50%	96.92%
F1-score	97.35%	96.89%

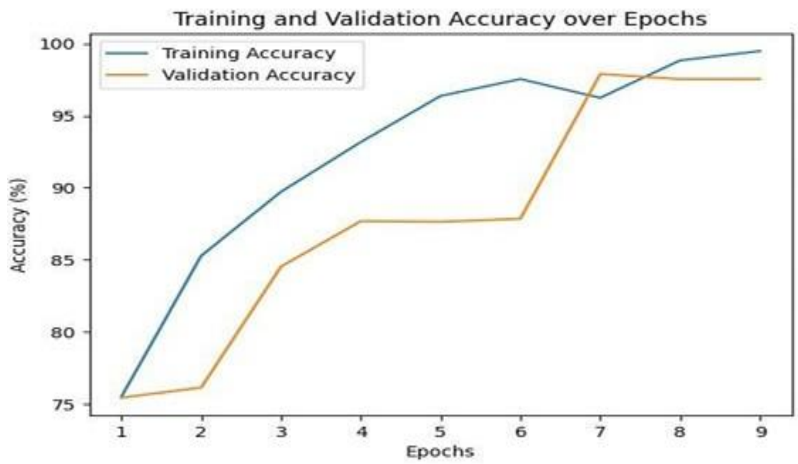


FIGURE 9. Training and validation accuracy over epochs

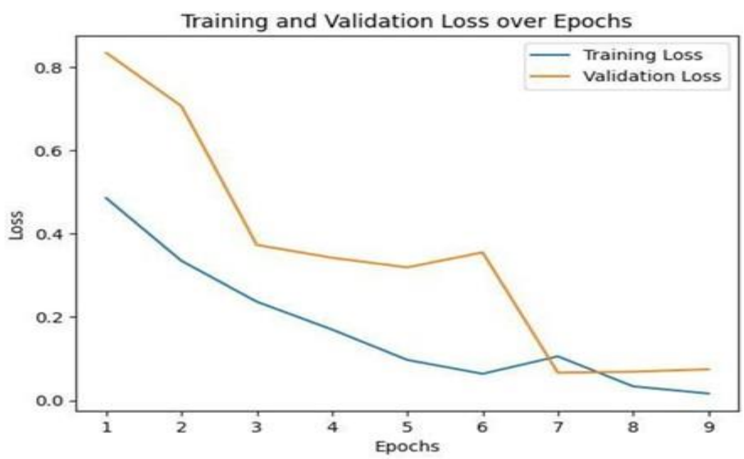


FIGURE 10. Training and validation loss over epochs

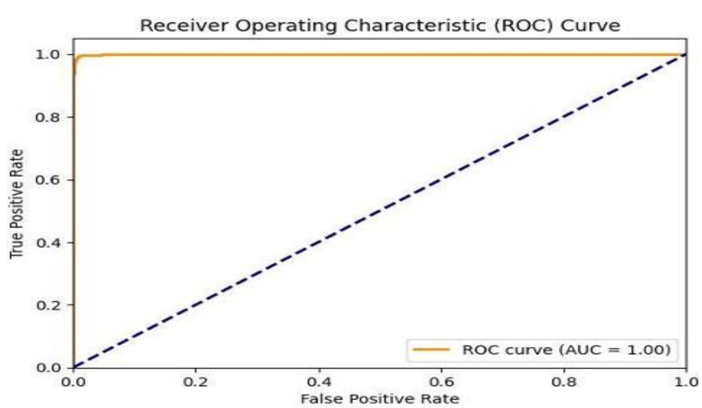


FIGURE 11. ROC curve

7. CONCLUSION

This research presented a high-accuracy toxic plant classification model using EfficientNet-B7, achieving superior performance compared to traditional approaches. The model has been integrated into a web-based interface, enabling real- time toxic plant identification for improved public safety and environmental awareness. Despite its strong performance, future work will focus on enhancing the model's robustness by expanding the dataset, incorporating additional plant species, and leveraging advanced techniques such as attention mechanisms and self-supervised learning. These improvements aim to further reduce misclassification and enhance real-world applicability.

REFERENCES

[1] K. Alobeidli, M. Shatnawi and B. Almansoori, ”Image Classi- fication for Toxic and Non-Toxic Plants in UAE,” 2023 Ad- vances in Science and Engineering Technology International Con- ferences (ASET), Dubai, United Arab Emirates, 2023, pp. 1-5, doi: 10.1109/ASET56582.2023.10180875.

[2] E. Bhatt and C. Greenberg, "Plant Toxicity Classification By Images," Journal of student research, vol. 12, no. 1, Feb 2023, doi: 10.47611/js- rhs.v12i1.4030. .

[3] Department of Health, Zayed Complex for Herbal Research and Traditional Medicine "Toxic Plants in UAE" available online: https://www.doh.gov.ae/en/resources/Supplement-Product.

[4] J. Huixian, ”The analysis of plants image recognition based on deep learning and artificial neural network,” IEEE Access, vol. 8, pp. 68828 -68841, 2020.

[5] Y. Sun, Y. Liu, G. Wang and H. Zhang, ”Deep Learning for Plant Identification in Natural Environment,” Computational Intelligence and Neuroscience, vol. 2017, 2017.

[6] Tufail, I. Ullah, R. Khan, L. Ali, A. Yousaf, A. U. Rehman, W. Alhakami, H. Hamam, O. Cheikhrouhou and Y.-K. Ma, ”Recognition of Ziziphus lotus through Aerial Imaging and Deep Transfer Learning Approach,” Mobile Information Systems, vol. 2021, 2021.

[7] T. H. Noor, A. Noor and M. Elmezain, ”Poisonous Plants Species Prediction Using a Convolutional Neural Network and Support Vector Machine Hybrid Model,” Electronics, vol. 11, no. 22, p. 3690, 2022.

[8] Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improve- ment. arXiv preprint arXiv:1804.02767.

[9] Manoranjan Dash, Design of Finite Impulse Response Filters Using Evolutionary Techniques - An Efficient Computation, ICTACT Journal on Communication Technology, March 2020, Volume: 11, Issue: 01

[10] Manoranjan Dash, "Modified VGG-16 model for COVID-19 chest X-ray images: optimal binary severity assessment," International Journal of Data Mining and Bioinformatics, vol. 1, no. 1, Jan. 2025, doi: 10.1504/ijdmb.2025.10065665.

[11] Manoranjan Dash et al.," Effective Automated Medical Image Segmentation Using Hybrid Computational Intelligence Technique", Blockchain and IoT Based Smart Healthcare Systems, Bentham Science Publishers, Pp. 174-182,2024

[12] Manoranjan Dash et al.," Detection of Psychological Stability Status Using Machine Learning Algorithms", International Conference on Intelligent Systems and Machine Learning, Springer Nature Switzerland, Pp.44-51, 2022.

[13] "Darknet: Open Source Neural Networks in C," Darknet: Open Source Neural Networks in C. https://pjreddie.com/darknet/ (accessed Dec. 11, 2022).

[14] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258).

[15] "GoogLeNet convolutional neural network - MATLAB googlenet," GoogLeNet convolutional neural network - MATLAB googlenet. https://www.mathworks.com/help/deeplearning/ref/googlenet.html (accessed Dec. 12, 2022).

[16] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.

[17] "Residual Networks (ResNet) - Deep Learning - GeeksforGeeks," GeeksforGeeks, https://www.geeksforgeeks.org/residual- networks- resnet-deeplearning/ (accessed Dec. 14, 2022).

[18] Ketwongsa, W., Boonlue, S., & Kokaew, U. (2022). A new deep learn- ing model for the classification of poisonous and edible mushrooms based on improved alexnet convolutional neural network. Applied Sciences, 12(7), 3409. https://doi.org/10.3390/app12073409

[19] Cho, J., Jeon, S., Song, S., Kim, S., Kim, D., Jeong, J., Choi, G., & Lee, S. (2019). Identification of toxic herbs using deep learning with focus on the sinomenium acutum, Aristolochiae man- shuriensis caulis, akebiae caulis. Applied Sciences, 9(24), 5456. https://doi.org/10.3390/app9245456.

[20] Manoranjan Dash, N.D. Londhe, S. Ghosh, et al., "Hybrid Seeker Optimization Algorithm-based Accurate Image Clustering for Automatic Psoriasis Lesion Detection", Artificial Intelligence for Healthcare (Taylor & Francis), 2022, ISBN: 9781003241409.

[21] Vytla, V., Ramakuri, S. K., Peddi, A., Srinivas, K. K., & Ragav, N. N. (2021, February). Mathematical models for predicting COVID-19 pandemic: a review. In Journal of Physics: Conference Series (Vol. 1797, No. 1, p. 012009). IOP Publishing).

[22] Ramakuri, S. K., Prasad, M., Sathiyanarayanan, M., Harika, K., Rohit, K., & Jaina, G. (2025). 6 Smart Paralysis. Smart Devices for Medical 4.0 Technologies, 112.