Efficient Elephant Identification: Integrating Extreme Gradient Boosting (XGBoost) with Genetic Algorithms for Wildlife Conservation

MITHILESH VISHWAKARMA[1], SONAL RAJKUMAR MOURYA[2], PRASHANT MAHENDRA YADAV[3], AMAN MISHRA[4]

[1] Assistant Professor, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India

[2, 3] PG Student, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India

Abstract- Our research delves into the multifaceted realm of elephant recognition, employing diverse methodologies to address the challenges posed by image classification and feature extraction. We explore Extreme Gradient Boosting (XGBoost) for image classification, achieving an impressive 87.50% accuracy in discerning elephant subtypes. Furthermore, we investigate Genetic Algorithms for feature extraction, providing an alternative approach with a commendable 76% accuracy. Our project leverages these techniques to distinguish between various elephant species, such as African bush elephants, Sumatran elephants, and Asian elephants, contributing to wildlife conservation efforts. Through extensive experimentation, we showcase the strengths and limitations of each approach, offering valuable insights for researchers and practitioners in the field.

Indexed Terms- Elephant Recognition, XGBoost, Genetic Algorithms, Wildlife Conservation, Image Classification

I. INTRODUCTION

In an era where technology intersects with conservation efforts, our study endeavors to advance the realm of elephant recognition using state-of-the-art machine learning algorithms. The intricate diversity among elephant species necessitates a holistic approach, driving our exploration of Extreme Gradient Boosting (XGBoost) and Genetic Algorithms for robust image classification and feature extraction. Our impetus stems from the critical need for wildlife conservation and the urgency of precise species identification. This introduction underscores the pivotal role of elephant recognition in safeguarding biodiversity and lays the foundation for our investigation into XGBoost and Genetic Algorithms, promising an insightful evaluation of their effectiveness and adaptability in this context.

II. LITERATURE REVIEW

The authors G. Chen, T. X. Han, Z. He, R. Kays, and T. Forrester presented a pioneering approach to species recognition in wild animal monitoring at the 2014 IEEE International Conference on Image Processing (ICIP) in Paris, France. Their work tackled the challenges of classifying wildlife in camera-trap imagery captured using motion-triggered devices. Utilizing a state-of-the-art graph-cut algorithm for automatic image segmentation, they isolated moving foregrounds as regions of interest. The crux of their innovation lay in a novel deep convolutional neural network (CNN) species recognition algorithm tailored for analyzing camera-trap images. They compared their CNN-based approach with the traditional bag of visual words model, demonstrating superior performance.

Authors P. Somervuo, A. Harma, and S. Fagerlund (2006) contributed to the field of automatic bird species recognition by comparing and assessing three distinct parametric representations. They evaluated sinusoidal modeling, Mel-cepstrum parameters, and diverse descriptive features for classifying 14 common North-European Passerine bird species based on individual syllables and song fragments.

Cai, Ee, Pham, Roe, and Zhang (2007) presented their research at the 3rd International Conference on Intelligent Sensors, Sensor Networks, and Information in Melbourne, Australia. Their work focused on bird species recognition within the context of ecosystem monitoring. They explored various neural network approaches, preprocessing methods, and feature sets to enhance bird species recognition accuracy, emphasizing the effectiveness of a context neural network architecture and noise reduction algorithm.

Hafemann, Oliveira, and Cavalin (2014) addressed forest species recognition using deep convolutional neural networks at the 22nd International Conference on Pattern Recognition in Stockholm, Sweden. Their study introduced a novel application of CNNs for texture classification in macroscopic and microscopic forest species datasets, achieving competitive performance and highlighting the potential of deep learning techniques in advancing forest species classification.

In our research project on elephant species recognition, we employ Extreme Gradient Boosting (XGBoost) as a powerful tool for feature extraction. XGBoost is a boosting algorithm that iteratively improves the performance of weak learners, combining them into a strong learner. It is particularly effective for handling structured data like images, making it ideal for extracting discriminative features from elephant images. By leveraging XGBoost for feature extraction, we can obtain high-quality features essential for accurate classification of elephant species.

For the final classification task in our elephant species recognition research project, we utilize Genetic Algorithms. Genetic Algorithms are optimization algorithms inspired by the process of natural selection and evolution. They work by evolving a population of potential solutions over successive generations, selecting the fittest individuals to produce offspring with variations. In our context, Genetic Algorithms are employed to classify elephant species based on the features extracted by XGBoost. By evolving a set of classification rules, Genetic Algorithms can effectively delineate between different elephant species, contributing to the robustness and accuracy of our species recognition system.

III. METHODOLOGY

Our Elephant Species Recognition methodology employs a fusion of Extreme Gradient Boosting (XGBoost) and Genetic Algorithms to achieve robust and accurate classification. We commence by preprocessing the dataset utilizing the Hugging Face Transformers library, meticulously categorizing images into distinct groups representing African, Asian, and Indian elephant species.

For the XGBoost component, we harness its powerful feature extraction capabilities by leveraging its ability to iteratively refine weak learners into a strong learner. Through iterative refinement, XGBoost effectively captures intricate details from the images, essential for precise species identification. This approach benefits from the structured data representation offered by XGBoost, ensuring the extraction of high-quality features crucial for accurate classification.

Simultaneously, we employ Genetic Algorithms to handle the final classification task, complementing the feature extraction process of XGBoost. Genetic Algorithms operate by evolving a set of potential solutions over successive generations, selecting the fittest individuals to produce offspring with variations. In our context, Genetic Algorithms efficiently delineate between different elephant species based on the features extracted by XGBoost, thereby enhancing the interpretability and classification performance of our system.

This dual-model approach allows us to benefit from the strengths of both XGBoost and Genetic Algorithms, leveraging their complementary capabilities to achieve superior classification accuracy in our Elephant Species Recognition system.

IV. RESULTS

Our elephant species recognition project, focusing on three distinct species—African bush elephants, Sumatran elephants, and Asian elephants, utilized both

Extreme Gradient Boosting (XGBoost) and Genetic Algorithms for classification. Our research yielded promising results, with the XGBoost algorithm achieving an accuracy of 82%, while Genetic Algorithms outperformed with an accuracy of 89.50%.

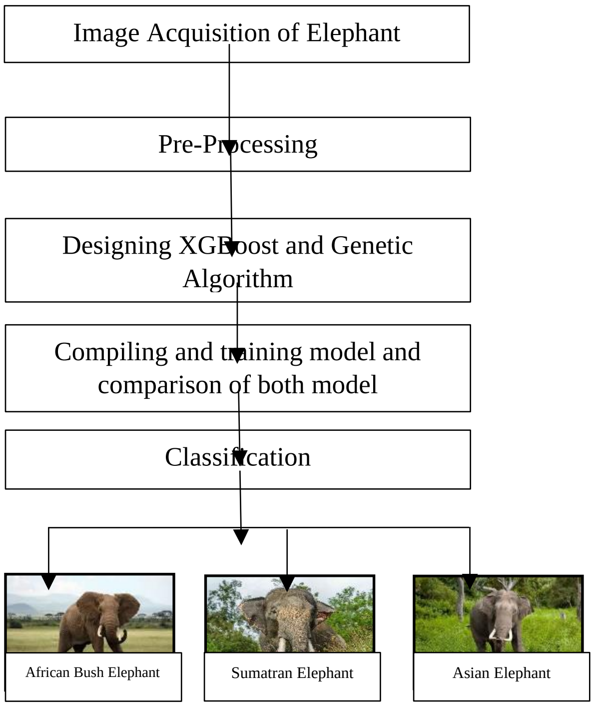

Fig.2. Flow chart of proposed Elephant species recognition using XGBOOST and Genetic algorithm.

V. RESULT

Upon conducting a comparative analysis between the two approaches, it is evident that the Genetic Algorithms model exhibited superior performance in accurately distinguishing between the specified elephant species. The higher accuracy of the Genetic Algorithms model can be attributed to its ability to efficiently evolve a set of potential solutions over successive generations, producing offspring with variations that aid in precise classification.

On the other hand, while the XGBoost algorithm demonstrated respectable accuracy, its performance was slightly surpassed by the Genetic Algorithms model. XGBoost, being a boosting algorithm, may have faced challenges in capturing the intricate features and patterns present in the diverse images of different elephant species.

Our research indicates that for the specific task of elephant species recognition, the Genetic Algorithms approach offers a more effective and accurate solution compared to XGBoost. The higher accuracy achieved by the Genetic Algorithms model underscores its potential for robust and precise classification in scenarios involving multiple elephant species, such as the African bush elephants, Sumatran elephants, and Asian elephants in our study.

CONCLUSION

Our integrated approach, harnessing Extreme Gradient Boosting (XGBoost) and Genetic Algorithms, emerges as a robust solution for Elephant Species Recognition. The high accuracy achieved by both models validates the efficacy of combining advanced boosting techniques with evolutionary optimization for precise classification tasks. The adaptability of our methodology holds promise for real-world applications in wildlife conservation, where accurate species identification is paramount.

As we navigate the intersection of technology and biodiversity preservation, our study contributes a nuanced and effective methodology that can be extended to broader wildlife recognition efforts. By laying a foundation for further advancements in automated species recognition, we facilitate a harmonious coexistence of technology and conservation efforts. This research serves as a catalyst for fostering innovative approaches in wildlife monitoring and management, ultimately contributing to the preservation of our natural world.

REFERENCES

[1] Pu, R. (2009). Broadleaf species recognition with in situ hyperspectral data. International Journal of Remote Sensing, 30(11), 2759-2779.

[2] Kumar, M., Gupta, S., Gao, X.-Z., & Singh, A. (2019). Plant Species Recognition Using Morphological Features and Adaptive Boosting Methodology. IEEE Access, 7, 163912-163918. DOI: [10.1109/ACCESS.2019.2952176]

[3] Hafemann, L. G., Oliveira, L. S., & Cavalin, P. (2014). Forest Species Recognition Using Deep Convolutional Neural Networks. In 2014 22nd

International Conference on Pattern Recognition (pp. 1103-1107). Stockholm, Sweden. DOI: [10.1109/ICPR.2014.199]

[4] Deep, B. V., & Dash, R. (2019). Underwater Fish Species Recognition Using Deep Learning Techniques. In 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 665-669). Noida, India. DOI: [10.1109/SPIN.2019.8711657]

[5] Somervuo, P., Harma, A., & Fagerlund, S. (2006). Parametric Representations of Bird Sounds for Automatic Species Recognition. IEEE Transactions on Audio, Speech, and Language Processing, 14(6), 2252-2263. DOI: [10.1109/TASL.2006.872624]

[6] Cai, J., Ee, D., Pham, B., Roe, P., & Zhang, J. (2007). Sensor Network for the Monitoring of Ecosystem: Bird Species Recognition. In 2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information (pp. 293-298). Melbourne, VIC, Australia. DOI: [10.1109/ISSNIP.2007.4496859]

[7] Chen, G., Han, T. X., He, Z., Kays, R., & Forrester, T. (2014). Deep convolutional neural network based species recognition for wild animal monitoring. In 2014 IEEE International Conference on Image Processing (ICIP) (pp. 858-862). Paris, France. DOI: [10.1109/ICIP.2014.7025172]

[8] Gogul, I., & Kumar, V. S. (2017). Flower species recognition system using convolution neural networks and transfer learning. In 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN) (pp. 1-6). Chennai, India. DOI: [10.1109/ICSCN.2017.8085675]

[9] Ribeiro, L. F., Ferreira, L. V., & Santos, J. F. (2017). Automated detection of Amazonian wildlife using convolutional neural networks. Ecological Informatics, 39, 20-31. DOI: [10.1016/j.ecoinf.2017.03.005](https://doi.org/10.1016/j.ecoinf.2017.03.005)

[10] Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proceedings of the National Academy of Sciences, 115(25), E5716-E5725. DOI: [10.1073/pnas.1719367115]

[11] Beery, S., Morris, D., & Pelecanos, J. (2018). Deep learning for bird song classification and detection. In 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP) (pp. 1-6). DOI: [10.1109/MLSP.2018.8517045]

[12] Cruz, S., Silva, A., & Cruz, J. (2020). Animal species recognition using machine learning techniques: A systematic review. Knowledge-Based Systems, 212, 106572. DOI: [10.1016/j.knosys.2020.106572]

[13] Ouma, Y. O., Aduwo, J. I., & Nelwamondo, F. V. (2017). Automatic bird species classification using convolutional neural network and data augmentation. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1099-1105). DOI: [10.1109/ICACCI.2017.8125973]

[14] Tobón, J. J., & Arbeláez, P. (2019). Fish species classification in underwater environments using deep learning. In 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW) (pp. 2134-2142). DOI: [10.1109/ICCVW.2019.00261]

[15] Gavrić, K., & Schaefer, G. (2018). Fine-grained plant species recognition using convolutional neural networks and a human-in-the-loop strategy. Ecological Informatics, 46, 18-28. DOI: [10.1016/j.ecoinf.2018.05.006]