Snake Species Classification Using ConvNeXtXLarge

Dr. Jayavrinda Vrindavanam[1], Pradeep Kumar[2], Chandrashekar N[3]

[1,2,3]Computer Science and Engineering (Artificial Intelligence and Machine Learning), Dayananda Sagar University, Bangalore, India

ABSTRACT
In this paper we are studying present a novel approach to snake species classification, utilizing ConvNeXtXL, a cutting-edge convolutional neural network architecture, and integrating a Language Model like GPT-3.5 for detailed information retrieval. The methodology involves collecting a diverse dataset of snake images, preprocessing them for uniformity, and fine-tuning the ConvNeXtXL model to classify 80 different snake species efficiently. Additionally, GPT-3.5 generates informative textual descriptions about the classified images, enriching the dataset with contextual knowledge about snake behavior and ecology. To make the classification system accessible, a web application is developed using Streamlit, enabling users to upload snake images and receive both visual classification results and textual descriptions. This combined approach enhances species identification, facilitates research and conservation efforts, and promotes public engagement in snake biodiversity and conservation.
General Terms: Snake species classification, deep learning approach, transformer

Keywords: Deep learning, ConvNeXtXLarge, Image processing, LLM, transfer learning

1. INTRODUCTION

In world there are 3971 snake species out of which 600 are venomous species in the world and about 200 are able to kill a human. This is an overview of the snakes that pose a significant health risk to human, through snakebites or other physical trauma. It becomes very important for people how got bit by a snake to classify the species of the snake so that they can get the right treatment for the snake bit to avoid any physical trauma.

As per world health organization in world there are 81,00 to 138,00 people who dies each year from snake bites worldwide. WHO also estimated that 5.4 million people worldwide are bitten by snakes every year and 1.8 to 2.7 million cases of envenoming occur.

Bites by venomous snakes can cause paralysis that many prevent breathing, bleeding disorders that can lead to a fatal hemorrhage, irreversible kidney failure and tissue damage that can cause permanent disability and limb amputation. Agricultural workers and children are the most affected.

2. LITERATURE SURVEY

The research that has being conducted a lot more in resent time where they use concept of deep learning model based on CNN and yolo based model to classify class of snake. In this paper they have classified venomous snake species using SnakeCLEF2023 dataset where they have achieved and accuracy of 93.65% which is combination of F1 and other parameters. They have used the following model to get the required accuracy like ResNet50, ConnvNextL models to get this accuracy [2]. Similarly in this paper they have used image data of snakes to classify if an image of the snake is venomous and non-venomous insect and snakes using multi-model approach and performed inference using ensemble method the model used in the ensemble like MoblieNet, InceptionResNet and DenseNet using the weight of these trained model have performed inferencing using ensemble method to get an overall accuracy of 89% [3]. Other paper has made an repellent device and detection using raspberry pi where they have used Yolov5 model and the model achieved an precision of 99.5% accuracy [4]. Similarly, a lot of research are being conducted where the paper has used EfficientNet, DEIT, DinoV2 models where there were able to achieve an train loss of 0.03,0.46, 2.04 respectively [5]. There another paper that which does a classification on snake bit for a specific class of snake which Egyptian cobra the model that was used to perform the classification of snake bit using VGG16 model with an accuracy of 96.7% [6]. Similarly another paper used deep learning method to classify snake species where they have used model like VGG16, DenseNet121 and MobileNetV2 where VGG16 achieved the highest accuracy of 97.09% [7]. This paper have prosed an methodology using Yolo algorithm with an 100% representation of snake in bounding box [8]. Another paper have proposed a method using Yolo model which achieved an a precision percentage of 87% [9]. Another paper have proposed an a method using InceptionResNet which got an accuracy of 82.64%, was the highest in this paper but there other model processed in this paper are as follows ReseNet34 with an accuracy if 76.79% Xecption with an accuracy of 76.79% and InceptionV4 with an accuracy of 63.57% [11].Another paper used model like Yolov5 and R-CNN where they made a system which will detect snake and send an sms to people who are using the device with the name of the snake [12]. Other paper used CNN and InceptionV3 to achieve an accuracy of 90% for 2 snake classes [13]. Similarly, another paper has proposed an approach using object detection with some approach like BME_TMIT, CMP, FHDO_BCSG, SSN, SSN_MLRG from comparing the result from all this the paper achieved an Macro f1 score of 0.903 [14]. Another paper has used Yolov4 to detect snake and a blurring system to prevent unexpected appearances of snake and the model that they trained achieved an f1 score of 92% [15]. Another paper have proposed a model which classifies 6 classes of snake and they have used a lot of machine learning and deep learning models and the most powerful model that they have found out for these 6 class classification is MobileNetv2 which achieved an accuracy of 93.16% [15]. Another paper has proposed a snake classification using ResNetXt50-v2 and other model where used to perform the ensemble model which had an accuracy of 86% [16]. Similarly, another paper has proposed an approach to classify 5 classes of snake and used 3 deep learning model which are based on CNN which are DensNet, Vgg16 and MobileNet and achieved an validation accuracy of 78%, 17.2% and 62.7% [17].

3. METHODOLOGY

In this paper we are proposing a methodology using one of the largest deep learning models to classify 80 classes of snake where the images and the data was collected through web scraping the data. Data preprocessing techniques were applied to the image to extract more feature where the model is deployed in real world. We have also used concept of LLM to get relevant information based on the class of snake being classified by the deep learning model. In fig 1 image describes the approach used to solve the problem

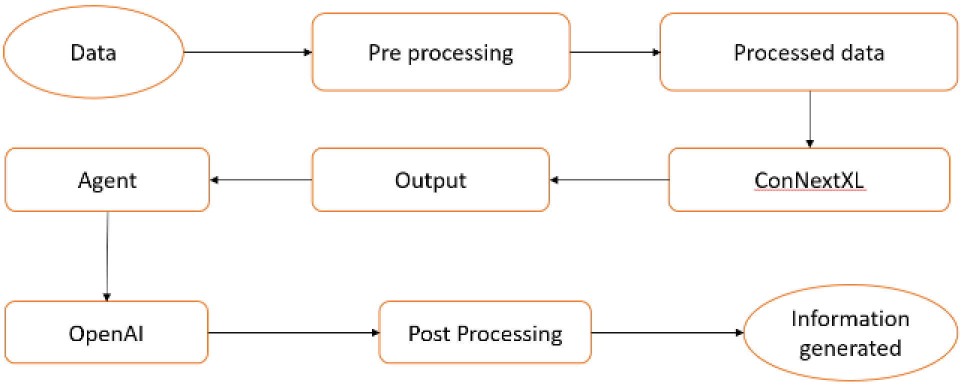

Figure 1 Methodology flow diagram

3.1 Data

The data used in this to get the required result and accuracy from the model so we have performed web scraping to collect image for 80 different classes of snake and performed some argumentation methods to provide the model with a diverse range of images so that it can classify the image which is captured in real time to avoid any miss classification across 80 classes of snake.

3.2 Data Pre-processing

The data pre-processing methods used to get real time simulation of camera module present in the farm fields are as follows

- Horizontal flipping: This data argumentation method performs horizontal flip on the input image. This mirrors the image along the horizontal axis respectively creating new variation of the original image.
- Vertical flipping: This data argumentation method performs vertical flip on the input image. This mirrors the image along the vertical axis respectively creating new variation of the original image.
- Rotation: A random rotation angle between 0 to 360 degree is applied to the image. This rotates the image clockwise by the specified angle introducing variation in orientation.
- Gaussian blur: It is a popular method for smoothing images by applying a convolution with a gaussian kernel. The script randomly selects a kernel size of either 3x3 or 5x5 pixels and applied gaussian blur to the image. This helps in simulating the effect of slight blurriness or noise in real world image.

3.3 ConvNeXtXLarge

This is a largest and powerful image recognition model based on transformer architecture. It was developed by DeepMind and achieved state-of-the-art result on several image classification benchmarks, including ImageNet-1k, ImageNet-21k and CIFAR-100. This model has the following information associated to the model layers. [10]

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

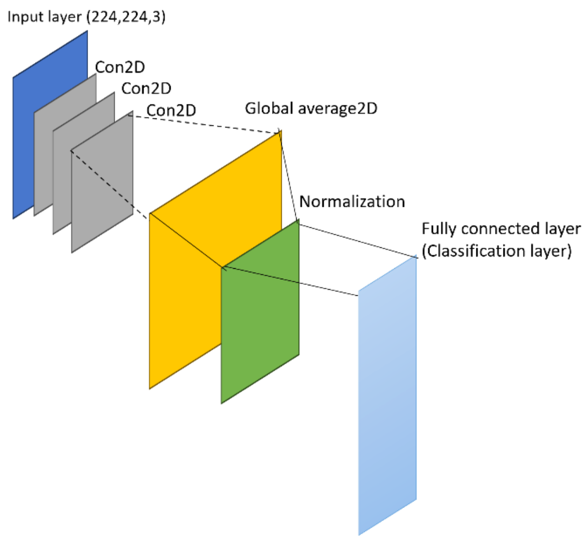

Figure 2 Architecture of ConNextXLarge

- Encoder: As this model is based on a transformer architecture where in this layer it will capture the relationship between image patches, represented by using matrix multiplication and SoftMax function.
- Decoder: As the masked self-attention to generate the final image representation, incorporating decoder only attention and encoder-decoder attention with masking mechanisms.
- ConvNextXL uses residual connection (F(x)+x) and skip connection to address vanishing gradients in deep networks. These involve simple summations of input and output values at different depths.
- For improved efficiency the model employes depth wise separable convolutions, factorizing the convolution operation into depth wise and pointwise convolution each using element-wise multiplication and summation.

3.4 Agent

Agent is an AI system that uses an LLM as its core but goes beyond simple text generation. Agent in LLM is where we provide the agent with tools which is used to perform operation on the LLM.

- The agent can go beyond just text by using tools and APIs to access information complete task, and even make decision based on context.
- Reasoning and planning can process information draw conclusion and plan their actions to achieve specific goals
- Autonomy where agent can adapt to situations without constant human intervention.

3.5 Large Language Model

Large language model is a type of AI where it has the capability in understanding and generating text. They play a significant role in the evolution of AI, offering innovative solution and raising various considerations. [1]

So here we have used LLM to generated information for the classified class of snake by the deep learning model and uses specific prompting method to get the required information from the LLM and the LLM model used is GPT-3.5. The tools like Wikipedia is used to perform RAG where it retrieves information from Wikipedia tools and generated information based on the information retrieved by the agent using tools is the processed to extract information based on the users prompt.

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

3.6 Post processing

Post-processing, in the context of information extraction, refers to the steps taken to refine and extract specific data or insights from raw output generated by a system. One common method used in post-processing is a string output parser, which involves analyzing text outputs to identify and extract relevant information based on predefined patterns or rules. This parsing technique essentially involves breaking down the output into smaller components, such as words or phrases, and then interpreting them to extract the desired data. For instance, in natural language processing tasks, a string output parser may be used to extract entities like names, dates, or locations from a piece of text. This process often involves a combination of techniques such as regular expressions, pattern matching, and linguistic analysis to accurately identify and extract the required information. Overall, string output parsing plays a crucial role in transforming raw outputs into structured and usable data, enabling further analysis or action based on the extracted information.

4. Results

The model has achieved an accuracy of 53% across 80 classes of snake. The below table shows the parameters regarding the model accuracy with respect to train, validation, test respectively.

Table 1 Result of the model

Model	Train accuracy (%)	Val accuracy (%)	Test accuracy (%)
ConNextXL	53	50	50

The below images Fig2 and Fig3 shows the output of the model and the information generated by LLM when the LLM is prompted to information about the classified snake species by the deep learning model.

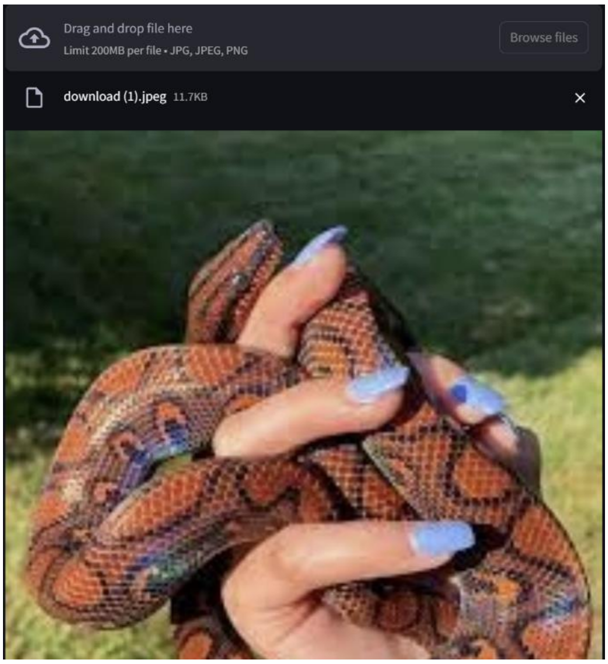

Figure 3 Analyze Serpent

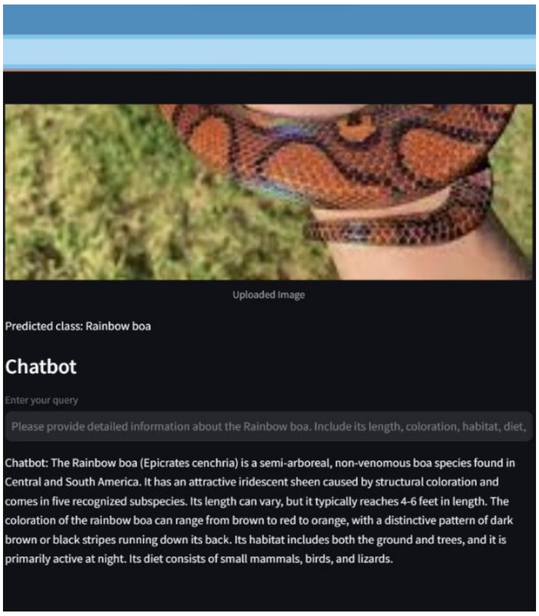

Figure 4 Analysis done using ConvNextXL Model and GPT-3.5

5. CONCLUSION

In conclusion, this study underscores the integration of emerging artificial intelligence methods, such as Language Model like LLM, to interface with results generated by deep learning models. Despite achieving a training model accuracy of 53% across 80 classes, the utilization of LLM significantly enhances the interpretability and contextual understanding of the classification outcomes. By employing agents to retrieve supplementary information from the web, the system enriches the analysis with additional insights, thereby advancing the efficacy and applicability of the overall approach. This synergy between advanced AI techniques holds promise for further advancements in enhancing the depth and breadth of automated classification systems, paving the way for more informed decision-making and comprehensive understanding in diverse domains.

6. Future Scope

The project aims to expand its scope by including more snake species, thus improving biodiversity understanding and conservation efforts. It plans to integrate newer deep learning models to stay updated with technological advancements. Additionally, the project envisions using IoT devices for real-time snake species identification on farms, aiding in wildlife management decisions. This approach reflects the project's dedication to using technology to address real-world agricultural challenges and promote sustainable coexistence between humans and wildlife.

In summary, the project seeks to broaden its impact by including more snake species and integrating advanced deep learning models. It also plans to use IoT devices for real-time species identification on farms, highlighting its commitment to addressing agricultural challenges while promoting sustainability.

7. ACKNOWLEDGMENTS

We are deeply grateful to Dayananda Sagar University for their invaluable support and provision of resources, which have played a pivotal role in the successful culmination of this research endeavor. Furthermore, we would like to extend our heartfelt appreciation to our mentor, Prof. Pradeep Kumar, for their steadfast guidance, unwavering support, and profound expertise that have been indispensable throughout this enriching research journey.

8. REFERENCES

1. Nguyen-Mau, Toan, et al. "An information fusion based approach to context-based fine-tuning of GPT models." *Information Fusion* 104 (2024): 102202.
2. Hu, Feiran, et al. "Watch out venomous snake species: A solution to snakeclef2023." *arXiv preprint arXiv:2307.09748* (2023).
3. Samarasinghe, Buddhin Saroj, and K. B. N. Lakmali. "A Deep Ensemble Learning Approach for Venomous & Non-Venomous Serpents and Insect Species Identification." *2023 3rd Asian Conference on Innovation in Technology (ASIANCON)*. IEEE, 2023.
4. Gupta, Vaishnavi, et al. "Design and Implementation of Snake Detection Cum Repellent Device." *2023 International Conference on IoT, Communication and Automation Technology (ICICAT)*. IEEE, 2023.
5. Balana, Aarti. "Metric Weighted Ensemble Focal Loss for Snake Species Identification." (2023).
6. Hassan, Ahmed, Mohamed Elhoseny, and Mohammed Kayed. "Hierarchical cloud architecture for identifying the bite of “Egyptian cobra” based on deep learning and quantum particle swarm optimization." *Scientific Reports* 13.1 (2023): 5250.
7. Ahmed, Kareem, Mai A. Gad, and Amal Elsayed Aboutabl. "Snake species classification using deep learning techniques." *Multimedia Tools and Applications* (2023): 1-42.
8. Naresh, E., et al. "A Novel Framework for Detection of Harmful Snakes Using YOLO Algorithm." *SN Computer Science* 5.1 (2023): 52.
9. Cardoso, Emerson Álvaro, Felermino DMA Ali, and Saide Manuel Saide. "Snake Species Identification Using Deep Convolutional Neural Networks." *2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2022.
10. Liu, Zhuang, et al. "A convnet for the 2020s." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.
11. Deshan, P. D. R., et al. "Smart snake identification system using video processing." *TENCON 2021-2021 IEEE Region 10 Conference (TENCON)*. IEEE, 2021.
12. Othman, Zainab, et al. "Snake Species Identification using Digital Image Processing." *2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*. Vol. 6. IEEE, 2021.
13. Picek, Lukáš, et al. "Overview of SnakeCLEF 2021: Automatic snake species identification with country-level focus." (2021).
14. TEKGÜL, Selay, and Gökçe NUR YILMAZ. "Snake detection and blurring system to prevent unexpected appearance of snake images on visual media sources using deep learning." *Journal of Artificial Intelligence and Data Science* 1.2 (2021): 125-135.
15. Rajabizadeh, Mahdi, and Mansoor Rezghi. "A comparative study on image-based snake identification using machine learning." *Scientific reports* 11.1 (2021): 19142.
16. Kalinathan, Lekshmi, et al. "Automatic Snake Classification using Deep Learning Algorithm." *CLEF (Working Notes)*. 2021.
17. Vasmatkar, Mrugendra, et al. "Snake species identification and recognition." *2020 IEEE Bombay Section Signature Conference (IBSSC)*. IEEE, 2020.