



A Lightweight AI System for Snake Species Identification and Emergency Medical Guidance

Zhang Enhong and Seung Hwan Kang*

Payap University, Chiang Mai, 50000, Thailand
seung_h@payap.ac.th

Abstract. This project aims to establish an automated snake species classification system that can reduce the dangers posed during snake bite incidence in rural areas. Even though it is essential to determine the snake species to react accordingly in time, this is a challenging task, especially for laymen. Therefore, this paper proposes a lightweight deep learning model, implemented using the Mobilenetv2 model, which is more conducive and well-balanced between complicity and accuracy on standard computer hardware. This experiment utilized a collection of images of five different snake species obtained from the Kaggle platform, which consists of more than 17,000 images. The proposed method is equipped with rigorous preprocessing of images, involving normalization, edge detection processes using OpenCV to focus on skin patterns of the snakes, and data augmentation approaches, including rotation and zooming techniques, resulting in significantly improved accuracy of the model. Even though the accuracy level was approximately 77% during the primary training process, it significantly improved to 90% after the finetuning and optimization processes of the model.

Keywords: Snake Species Identification, MobileNetV2, Image Classification.

1 Introduction

Snakebite envenomation is recognized by the World Health Organization (WHO) as a neglected tropical disease that continues to impact millions of people annually. In regions such as Asia, Africa, and Latin America, many communities reside near snake habitats. According to the WHO, approximately 5 million people are bitten each year, resulting in thousands of deaths and numerous long-term injuries [1]. These incidents affect rural areas where medical services are limited and travel times to hospitals can be extensive.

One of the primary challenges in snakebite treatment is the precise identification of the species involved. Because different snakes require specific types of antivenom, accurate identification is critical for clinical success. However, many species exhibit similar colors and patterns, and the chaotic nature of an accident often prevents victims from observing the snake clearly. Consequently, medical staff must often rely on vague descriptions, which increases the probability of incorrect treatment. Traditional methods, such as consulting printed guides or seeking remote experts, are typically too slow

for emergency scenarios. As most small clinics lack specialized herpetologists, there is a need for a more fast and accessible digital solution for species recognition [2].

The difficulty of species identification is a critical safety issue. Even when a photograph is available, non-experts and general practitioners in small clinics may struggle to distinguish between morphologically similar species. For instance, research has shown that in North America, the harmless Northern Watersnake is frequently mistaken for the venomous Copperhead due to similar banding [3]. Similar identification challenges occur across Asia with various banded species. Administering the incorrect antivenom is dangerous, as it may cause severe allergic reactions or worsen the patient's condition, while delaying treatment due to uncertainty remains life-threatening. Therefore, a reliable tool is needed that can classify snake species quickly and function on standard hardware or mobile devices, enabling faster decision-making for both medical workers and the public.

The motivation for this study stems from witnessing the fear and uncertainty that communities experience during snake encounters. In many rural areas across Thailand and Southeast Asia, immediate medical guidance is scarce, and reliable identification tools are rarely available. The potential for AI to reduce medical risks and save lives makes this project both meaningful and impactful. Unlike purely academic exercises, this research addresses a tangible public health challenge faced by communities worldwide. From a technical perspective, this project provides a robust framework for applying computer vision and machine learning. Developing the system requires integrating deep learning architectures like MobileNetV2 with complex image preprocessing to create a complete, functional tool for real-world application [4].

During the entire process of developing this system, many prominent hurdles stood out to be overcome. The disparities among online sources in such a way that light conditions are not similar in all images and images with improper labels needed rigorous processing to eliminate all disparities to form reliable models. Moreover, careful resource management was required to reduce batch size during model training procedures due to the implementation on hardware with lower computational capacity and RAM. Lastly, one prominent drawback in the present system is that it comprises images from only five identified species [5]. This represents a proven successful demonstration but points toward extending it to encompass images from many more species, especially from Thailand.

2 Literature Review

The need for rapid identification of species of snakes has long been established from the viewpoint of the efficacy of anti-venoms. Traditional methods of identification involve the use of image descriptions that can result in significant delays. Innovations that have occurred in the application of artificial intelligence concepts, particularly with regards to the application of computer vision technology, can prove to be instrumental in revolutionizing the aspects of individual identification.

Recently, research has set up a standard regarding the identification of different species of snakes across the world through deep learning. Durso et al. analyzed the performance of computer vision models on a collection of over 300,000 images, thereby proving that although similarities in visual patterns are an existing problem, computer vision models can be effective and even succeed human levels of performance in many parts of the world [3]. This study illustrates how effective these types of computer vision models are, especially while being trained on data sets that cover species variation.

Furthermore, the implementation of these models in clinical settings has been explored by Bolon et al. [2]. Their research focused on the practical application of AI in global health, emphasizing that the primary barrier to adoption is not just accuracy, but the ability of the model to function in low-resource environments [2]. Their findings suggest that for AI to be truly impactful in the "snakebite belt" of Asia and Africa, it must be deployable on hardware that does not require constant internet connectivity or high-end GPUs.

From a technical perspective, the evolution of lightweight neural networks has made field-based identification possible. MobileNetV2, introduced by Sandler et al., utilizes a unique "inverted residual" structure and linear bottlenecks to significantly reduce the number of parameters and computing costs without a substantial sacrifice in accuracy [4]. This architecture is designed specifically for mobile and edge devices. Beyond general classification, Rajabizadeh and Rezghi demonstrated that deep learning can successfully distinguish between venomous and non-venomous snakes by focusing on specific scale patterns, even when the image quality is suboptimal [6].

MobileNetV2 was selected for this project over heavier models like ResNet50 or VGG16 due to these practical considerations. While larger models can offer slightly higher accuracy, they require significant memory and computational power that are often unavailable in rural medical clinics or on standard portable laptops. MobileNetV2 offers a superior balance of speed and efficiency, making it the most suitable candidate for a real-time system designed to assist in emergency medical decision-making.

3 Methodology

This session explains the design, development, and testing of the artificial intelligence-based snake recognition system. This was done following various stages in the software development lifecycle and machine learning such as requirement analysis, design, data gathering, model creation, implementation, testing, and evaluation by use of the gathered data.

3.1 Requirement Analysis

Requirement analysis helps set the abilities and standards of the system. Functionally, it includes functionalities such as uploading images, classifying the species, as well as providing details such as venom information, geographic distribution, and emergency contact information. Other functionally related requirements include metrics such as the accuracy of prediction and the F1-score of the model.

Non-functionally speaking, the requirement aims at being user friendly, having at least an 80% accuracy rate within three seconds, and scalability in the future.

3.2 System Analysis and Design

With the requirements in place, the system design involved creating a structure that has four main layers: input layer, processing layer, model layer, and output layer. The input layer is responsible for receiving the uploaded images. The processing layer deals with transformations like resizing and normalization. The model layer is where the trained deep learning algorithm runs. The output layer displays the result in the form of the predicted species name and associated safe information to the user.

3.3 Data Collection and Preprocessing

Data gathering and preprocessing are key to the success of the proposed model. The data set is obtained from Kaggle and includes 17,389 images of five species of snakes, namely Black Rat Snake (3,453), Common Garter Snake (3,661), DeKay's Brown Snake (3,547), Northern Watersnake (3,087), and Western Diamondback Rattlesnake (3,641). Each category's folder is verified for quality and correctness. The data set is divided into 70 percent for training and testing, 20 percent for validation, and 10 percent testing.

Despite its huge size, some of the images presented wrongly labeled bodies or distracting environments consisting of leaves, rocks, and human hands that could lead to overfitting. In view of these concerns, a vetting process was conducted to remove unwanted images (See Fig. 1).



Fig. 1. Sample images of target snake species with varied lighting and backgrounds.

3.4 Model Architecture

MobileNetV2 was chosen as the preferred lightweight convolutional neural network (CNN) due to its effectiveness and efficiency. In contrast to baseline CNNs and the ResNet network, MobileNetV2 gave an accuracy of 77%, unlike the others, whose accuracy ranged between 65% and 72%. This is also attributed to the use of depth-wise separable convolutions in MobileNetV2, which reduced memory usage.

3.5 Training Configuration

The model was built and trained by employing the Adam optimizer and category cross-entropy loss. During the process of network training, there were a few callbacks. These

included ModelCheckpoint for storing and tracking the best network, a callback that automatically adjusted the learning rate because validation loss did not improve, and EarlyStopping to avoid overfitting. The accuracy was at a level of 77%, but this was improved through data augmentation, fine-tuning, and class weighting to about 90%.

3.6 Preprocessing Techniques

For the preprocessing of images, standardization to a uniform size of 224x224 was conducted, followed by scaling of the intensity range of the pixels to a range between 0 and 1. For improving the variety of images and avoiding overfitting, image augmentation was considered using rotation, scaling, shifting, and horizontal flipping (See Fig. 2).



Fig. 2. Data augmentation through rotation to improve model robustness.

Additionally, the OpenCV library for edge detection was incorporated in the image to highlight the patterns of scales and edges by zooming in on the body of the snake instead of the background (See Fig. 3).



Fig. 3. Edge detection to highlight snake morphological features and patterns.

3.7 User Interface

The deployment phase entailed the integration of the machine learning model with an interface created using Streamlit, a Python library used to create simplified web apps. The program enables the user to upload an image of the snake, which is then automatically resized and preprocessed before being fed into the MobileNetV2 model. The output parameters include the confidence level, venomous nature, and contact information of the virtual hospital for venomous snakes (See Fig. 4).

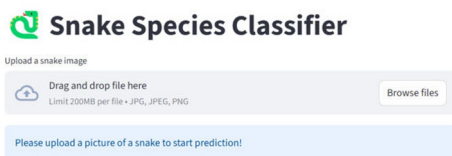


Fig. 4. Streamlit interface for species classification and medical guidance.

4 Experimental Result

Quantitative performance evaluation for the model was done with metrics such as accuracy, precision, recall, and F1-score. While accuracy can be defined as a measure of the general percentage of correct predictions, precision and recall are metrics showing the performance of a model on positive cases without false alarms or missed cases. The F1-score can be viewed as the harmonic means of precision and recall and is therefore a balanced metric for performance across all classes.

As illustrated in the performance summary, the model was effective across the five species. DeKay's Brown Snake and Western Diamondback Rattlesnake were the two most successful with a score of 0.84 and 0.86 F1-score, respectively. The Common Garter Snake followed at 0.79, then the Northern Watersnake at 0.76, and finally the Black Rat Snake proved to be the toughest for the model at 0.71 F1-scores (See Fig. 5).

	precision	recall	f1-score	support
black_rat_snake	0.73	0.68	0.71	346
common_garter_snake	0.72	0.87	0.79	367
dekays_brown_snake	0.92	0.77	0.84	356
northern_watersnake	0.70	0.82	0.76	310
western_diamondback_rattlesnake	0.93	0.80	0.86	365
accuracy			0.79	1744
macro avg	0.80	0.79	0.79	1744
weighted avg	0.80	0.79	0.79	1744

Fig. 5. Summary of evaluation metrics across five snake species.

The confusion matrix gives an accurate representation of how the classifier classifies the five species of snakes as the matrix represents the true predictions along with incorrect predictions. The Western Diamondback Rattlesnake is classified accurately at 80% with a precision of 93%. But, there is a lot of confusion with respect to the Black Rat Snake which gets misclassified 43 times and mostly confuses the common garter snake.

Confusion matrix:					
[	237	50	6	43	10]
[	21	321	12	11	2]
[	15	53	275	11	2]
[	27	16	5	255	7]
[	24	7	1	42	291]]

Fig. 6. Confusion matrix showing predicted versus actual species classification.

Testing the real-world application proved that the system was useful. When an image is uploaded into the Streamlit interface, the system can identify the image and give instant feedback. For example, when test examples were used, high levels of probability were obtained for classification of DeKay’s Brown Snake: 1.0, Western Diamondback Rattlesnakes: 0.96 and 0.87, and for Black Rat Snake: 0.98 (See Fig. 7).

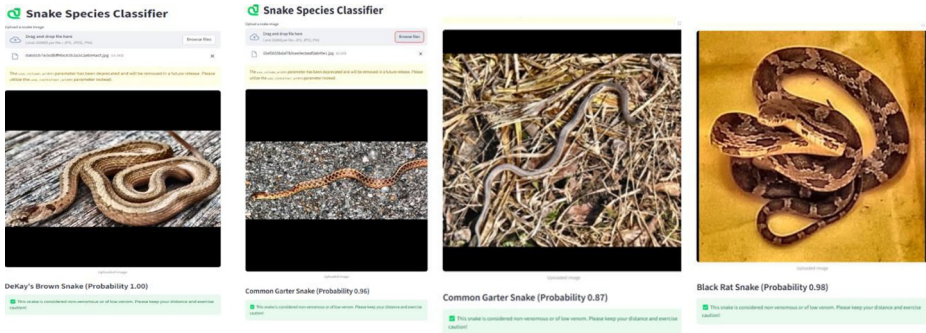


Fig. 7. Web interface test samples showing high-confidence classification results.

The application not only gave a high confidence score but also properly displayed the venomous and hospital contact details (See Fig. 8).

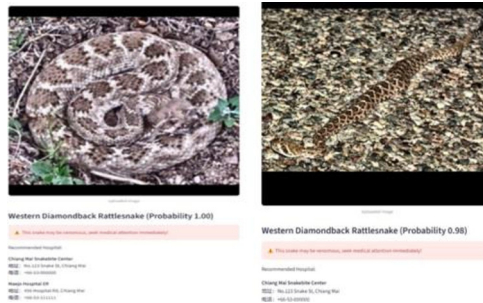


Fig. 8. Application output displaying venom status and localized hospital data.

These results demonstrate the efficacy of the union of the MobileNetV2 model and the user-friendly interface to effectively facilitate quick identification in the actual scene.

5 Conclusions

This project successfully reached its primary objective, which was the design and implementation of an AI-driven system that could identify snake species and provide critical information on their safety. With transfer learning and utilizing the architecture of MobileNetV2, the model significantly improved, starting with its baseline at 77% to about 90% at its best performance. A very important lesson learned in this research is that the development of artificial intelligence, the curation of the data, is as intense as the algorithmic implementation itself. Cleaning up the dataset and balancing the classes took plenty of time but was necessary to achieve reliable results. Another important aspect reviewed in this project is the potential of complex deep learning models to be pushed into practical, easily accessible web interfaces using the Streamlit framework.

Although the level of accuracy attained is surprisingly good, the study contains a few shortcomings that form the roadmap for the next level of development of the system. Currently, the system only works on the capability to detect the names of five species of snakes found in the region. For the system to be more functional and useful in the local environment, the next level of the system needs to be able to detect the names of many species found in Thailand and the Southeast Asia regions. Moreover, the system lacks the actual interaction interface for the contact number of the hospital, and it needs to be made functional so that the locations of the nearest hospitals that are authentic and possessing antivenom could be obtained from Google Maps services.

References

1. World Health Organization: Snakebite envenoming. <https://www.who.int/news-room/fact-sheets/detail/snakebite-envenoming>. (2025)
2. Bolon, I., Picek, L., Durso, A.M., Alcoba, G., Chappuis, F., Ruiz de Castañeda, R.: An artificial intelligence model to identify snakes from across the world: Opportunities and challenges for global health and herpetology. *PLOS Negl. Trop. Dis.* 16(8), e0010647. <https://doi.org/10.1371/journal.pntd.0010647> (2022)
3. Durso, A.M., Moorthy, G.K., Mohanty, S.P., Bolon, I., Salathé, M., Ruiz de Castañeda, R.: Supervised learning computer vision benchmark for snake species identification from photographs: Implications for herpetology and global health. *Front. Artif. Intell.* 4, 582110 <https://doi.org/10.3389/frai.2021.582110> (2021)
4. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4510–4520. IEEE, New York (2018)
5. Rahman, S.: Pre-processed snake images [Data set]. Kaggle. <https://www.kaggle.com/datasets/sameeharahman/preprocessed-snake-images> (2020)
6. Rajabizadeh, M., Rezghi, M.: A Comparative Study of Deep Learning Architectures for the Identification of Venomous Snakes. *Journal of Herpetology* 55(4), 384–392 (2021)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

