Deep Learning Approach for Snake Species Classification Using EfficientNet-B0, CNN, CNN-LSTM, and ANN Architectures

Dipta Paul[1], Nabila Tabassum[2], Taslima Azgar Esa[3]
[1] ID: 21225103315, 49(7) [2] ID: 21225103347, 49(7) [3] ID: 21225103332, 49(7)

***Abstract*—The classification of snake species is a problem because it is very difficult to distinguish the visual similarity between different species and the annotated data is also a matter of concern due to restricted quantification. This study fills the gap of a strong and automated solution through using and comparing various deep learning models, such as EfficientNet-B0, Convolutional Neural Networks (CNN), CNN-LSTM, and Artificial Neural Networks (ANN) to perform classification of the snake species based on image data. The relevance of this issue is the possibility of consequences in the field of ecological study, the health of the population, and the need to preserve biological diversity, as this will help make decisions with far-reaching implications. Even though the problem that focuses on SN classification with the pre-trained CNN has been considered in previous research, there are no comparative assessments of various architectures and optimization approaches. The paper addresses this gap by using transfer learning and extensive hyperparameterSearch on multiple models in a process to determine their usefulness. We have provided a deep learning pipeline comprising multi-model and performance benchmark of an architecture across models as well as a pool of experiments on a curated snake images dataset. The findings give promising results compared with conventional techniques and EfficientNet-B0 records the highest classification accuracy. These results indicate realm of usefulness with wildlife surveillance and ecological data analysis, and represent a scaleable and precise model to be utilized in real-world application both in the field of conservation and research.**

I. Introduction

In today's world, cats are more than just pets — they're beloved family members. Whether for veterinary care, adoption services, or even pet tech applications, being able to accurately identify a cat's breed can be incredibly useful. But doing that by eye isn't always easy. Many cat breeds look strikingly similar, and unless you're a trained expert, manual breed identification can be slow, inconsistent, and often unreliable.

That's where artificial intelligence (AI) comes in. With smart devices becoming more common and affordable, there's growing interest in using AI to solve practical problems — and identifying cat breeds from photos is a perfect example. Deep learning models, especially those that specialize in image recognition, are particularly well-suited for this kind of task.

However, most existing breed classifiers have some major limitations. Many are trained on a small number of breeds or can only tell if an image belongs to a certain breed or not (binary classification). That's not very helpful in real-world scenarios where there are dozens of possible breeds to choose from.

Our project set out to build a more capable and inclusive system — a multi-class image classifier that can distinguish between multiple cat breeds with high accuracy. We trained and evaluated two powerful deep learning models, ResNet50 and EfficientNet-B0, to see how well they perform when faced with the diversity and visual complexity of cat breed images. We also applied advanced techniques like image preprocessing, data augmentation, and Optuna-based hyperparameter tuning to get the best possible results.

Interestingly, EfficientNet-B0 not only achieved higher accuracy, but it also required less computing power — making it ideal for use in mobile apps or on lower-end devices. This opens the door to practical, real-time applications for shelters, pet owners, and veterinarians, even in resource-limited settings.

Of course, the project had its challenges. Some breeds look nearly identical, and others were underrepresented in the dataset, making training a balanced model difficult. We also had to work hard to keep the model efficient enough for real-world use, not just high-performance in lab conditions.

Despite these hurdles, our work shows that AI — especially deep learning — can be a powerful tool for reliable cat breed classification. In the next sections, we'll explain how the system was developed, explore the results, and share ideas for where this technology can go next.

Section II presents related works, Section III describes the proposed methodology, Section IV discusses results and analysis, and Section V concludes with future directions.

II. Related Works

The domain of computer vision Deep learning has transformed the domain of computer vision, especially in areas like image classification, pattern recognition and object detection. Automated snake species classification based on images is a new application that is of special interest to wildlife (conservation), scientific investigations, and medical applications (snakebite treatments). Because of minor differences in appearance between the species and the absence of huge labeled data, conventional taxonomies cannot always provide accurate answers. In this regard, the survey of related literature must be provided to get an idea of what methods are strong and weak and what models were best used and still there is a performance gap. Comprehensive literature review can assist in formulating existing issues and evaluating the importance of various deep learning frameworks, as well as explaining the choice of certain models (EfficientNet-B0, CNN, CNN-LSTM and ANN) to be used in this study. It also predetermines the possibility of thinking up new solutions and allows creating a more precise and generalizable baseline of classifying species of snakes and serpents. Progga at el. [1] developed an automated system to classify snakes as venomous or non-venomous using image data. The study employed a Convolutional Neural Network (CNN) for classification and achieved an accuracy of 91.30%. Despite the promising results, the research highlights challenges such as limited data availability, the need for further model tuning, and environmental variability affecting model performance.

Bolon at el. [2] proposed a system to support snakebite diagnosis through improved snake identification tools, aligning with the WHO's objective to reduce envenoming-related deaths by 50% by 2030. Using a snake photo dataset, the study applied a Vision Transformer model and achieved an accuracy of 94.2%. The work emphasizes the importance of addressing data gaps in snake ecology and snakebite epidemiology to further enhance diagnostic accuracy and model reliability.

Bloch at el. [3] focused on enhancing snake species classification by combining image and location data to improve care for snakebite patients. Utilizing the SnakeCLEF 2021 dataset, the study employed EfficientNets and Vision Transformers, achieving an accuracy of 82.88%. Challenges identified include the presence of irrelevant images in the dataset and time constraints during model training and evaluation.

Durso at el. [?] trained a model for accurate snake species identification using the Snake Species Identification Dataset. The study utilized EfficientNet and achieved a high accuracy of 96%. Despite the strong performance, the research notes issues such as species confusion and dataset bias, which can affect the model's generalizability in real-world scenarios.

Kareem Ahmed at el. [4] developed a deep learning-based fast and accurate system to recognize snake species from image data.The study utilized the SnakeCLEF2021 dataset and employed methodologies such as SOD and convolutional neural networks (CNNs), achieving an impressive accuracy of 99.32%.The research faced several challenges included lookalike species, age and location-based variations, cluttered backgrounds and limited data.

Mahdi Rajabizadeh at el. [5] aimed to enhance snake classification in Lar National Park using feature extraction and dimensionality reduction techniques. Using the Snake Images Dataset, methods like kNN, SVM, LR, LDA, and SVM were applied, achieving 93.16% accuracy. However the research faced several challenges such as pose variations, limited data, poor lighting, non-discriminative patterns, and limitations of museum specimens.

Lekshmi Kalinathan at el. [6] aimed to automate snake species identification using deep learning to reduce snakebites and support public health, conservation, and biodiversity. Using the SnakeCLEF2021 dataset, a ResNeXt50-V2 model was trained with Keras, achieving 85.7% accuracy. The research faced challenges including high species diversity, limited images, inaccurate bite victim reports and limited herpetological knowledge among healthcare providers.

Lukáš Picek at el. [7] proposed to create a platform for evaluating AI-driven snake species recognition and comparing it with human experts. Using the SnakeCLEF 2022 dataset, they applied vision transformers and ensemble methods and achieved 94.67% accuracy.The researc faced multiple challenges such as high within-species variation, incomplete species distribution, geographic data gaps, uneven data across species, and a long-tail issue in the dataset.

Cheng Zou at el. [8] presents a robust multimodal system for accurate snake species recognition with SnakeCLEF 2022 dataset. They applied Multimodal CNN-MLP-transformer model, logit adjustment, supervised and self-supervised learning achieving up to 82.72% accuracy.The study lacked class imbalance and to integrate complex multimodal data.

Rail Chamidullin at el. [9] proposes a deep learning approach using fine-tuned CNNs, species filtering, ensembling, and mixed precision training with SnakeCLEF2021.They applied fine-tuned ResNet-based models with optimized training strategies and country-specific prediction filtering achieving to 91.6% accuracy.The study lacked fine-grained classification, class imbalance, noisy data.

Anika Patel at el. [10] aims to develop a real-time AI platform using R-CNNs for accurate image-based identification of Galápagos snake species withGalapagos Snake species recognition dataset.They applied VGG16model achieving to 75% accuracy.The study lacked limited dataset, complex image backgrounds, frequent misclassifications, and poor model performance.

Karthik Desingu at el. [11] presents to build an automated snake species classifier using transfer learning with Snake species classification dataset.They applied ResNet V2, CNN models achieving to 42.96% accuracy.The study highlights challenges like Highly imbalanced,Limited labeled data,Limited computational resources.

Louise Bloch at el. [12] presents to improve snake species identification by combining object detection with Snake image dataset.They applied Mask R-CNN, EfficientNet models (B0 to B4) achieving to 59.4% accuracy.The study lacked Highly unbalanced class distribution,High intra-class variance and low inter-class variance.

Louise Bloch at el. [13] proposes To develop an automated system for snake species classification using deep learning and metadata with SnakeCLEF 2021.They applied Transformer (ViT-Large) models was trained, with predictions enhanced by incorporating binary country-level metadata achieving 87.13% accuracy. Handling the long-tailed class distribution and training on a large, imbalanced dataset with limited resources.

Mirunalini Palaniappan at el. [13] aims To develop a deep learning-based model for accurate snake species classification to aid rapid medical treatment with SnakeCLEF 2022.They applied VGG16, ResNet50, MobileNetV2, DenseNet121 achieving to 99.32% accuracy. Managing complex image backgrounds and imbalanced datasets with limited GPU resources.

LouiseBloch at el. [14] presents To enhance snake species identification using object detection, image classification, and geospatial data fusion with SnakeCLEF 2022.They applied EfficientNet/ConvNeXt achieving 78.085%. Major challenges included data imbalance, missing metadata, and high computational cost for large-scale model training.

III. Methodology

An end-to-end pipeline for image classification was developed. The input images were preprocessed and fed into two CNN models (ResNet50 and DenseNet-121). Outputs were compared using classification metrics.

A. Methodology Diagram

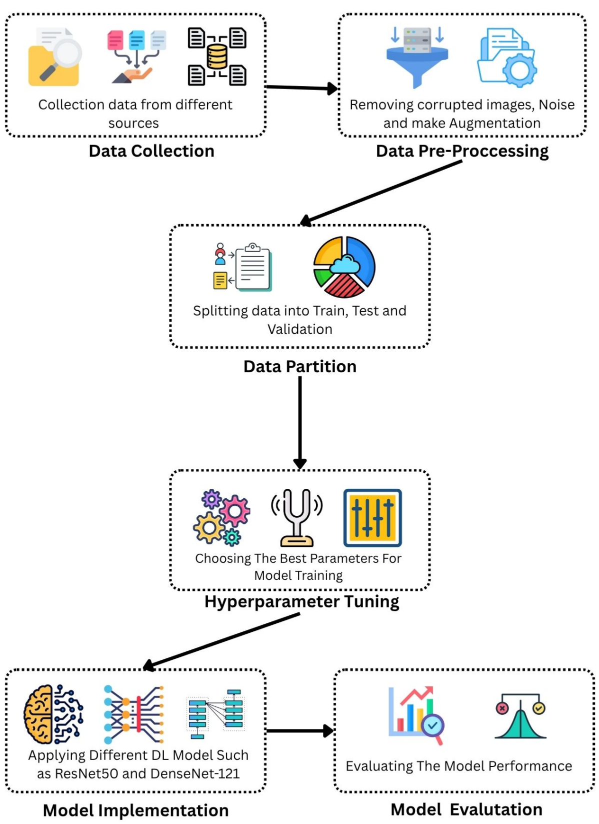


Fig. 1. Sample images from different classes in the dataset

B. System Architecture

The system consists of:

- **Input Layer:** Image ingestion
- **Preprocessing Module:** Resizing, normalization, augmentation
- **Feature Extractor:** ResNet50 / DenseNet-121
- **Classifier:** Fully connected layers
- **Output:** Predicted class label

C. Dataset and Preprocessing

We used a 67-class dataset sourced from Kaggle and selected 14 classes representing various cat breeds. The original dataset was imbalanced, with the number of images per class varying significantly—for example, the Bengal class had as few as 2,477 images, while others like Siamese had over 2,888. To address this imbalance, we applied the following preprocessing steps:

- Resized all images to 224×224 pixels
- Normalized pixel values
- Applied data augmentation (random flips, rotations) to increase samples and balance all classes to 3,500 images each
- Used weighted loss functions during training to further mitigate the impact of class imbalance

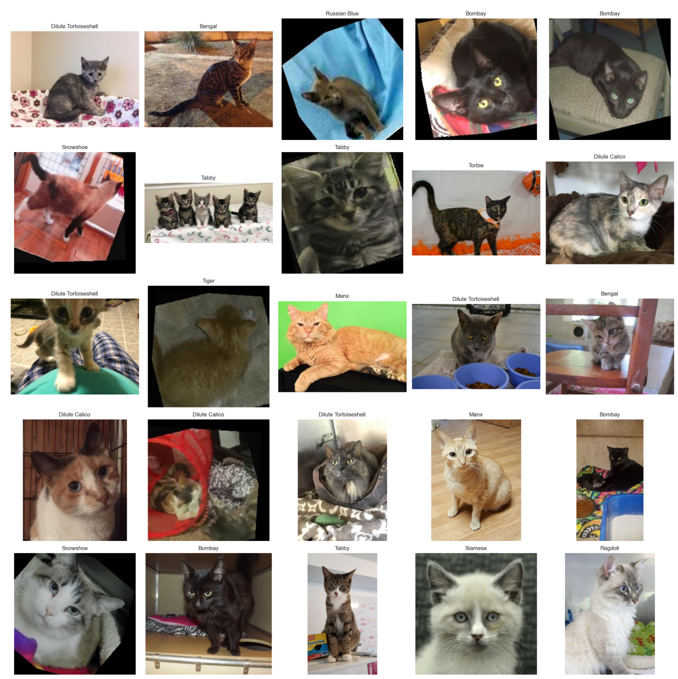


Fig. 2. Sample images from different classes in the dataset

D. Model Description

a) **ResNet50**
ResNet50 is a deep convolutional neural network with 50 layers that uses a technique called residual learning to address the vanishing gradient problem common in very deep models. It is built using a series of convolutional blocks that include shortcut (skip) connections, which enable the gradients to pass more easily through the network during training. This design improves learning and helps the model achieve high accuracy in image classification tasks while maintaining computational efficiency.

b) **DenseNet-121**
DenseNet121 is a convolutional neural network with 121

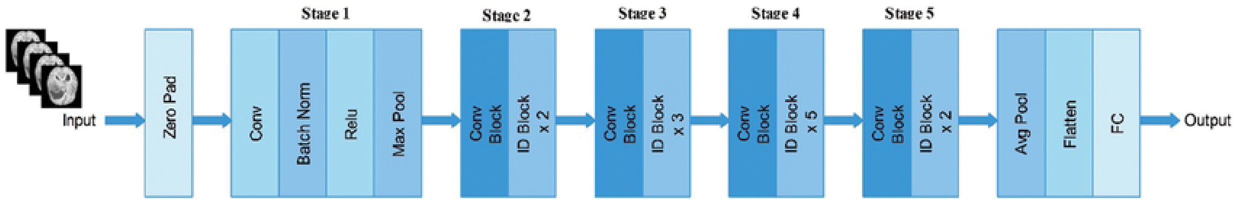


Fig. 3. Architecture of ResNet50

layers that uses dense connections to improve information flow between layers. In this architecture, each layer receives inputs from all previous layers and passes its own output to all subsequent layers within a dense block. This design reduces redundancy, encourages feature reuse, and helps prevent the vanishing gradient problem. DenseNet121 is known for its compact structure, efficient parameter usage, and strong performance in image classification tasks.

E. Implementation Details

The proposed models were implemented using the PyTorch deep learning framework. Two architectures—ResNet50 and DenseNet-121 were fine-tuned for the task of multi-cat breeds classification. Hyperparameter optimization was conducted using the Optuna framework to ensure optimal performance for each model.

For the DenseNet-121 model, the optimal hyperparameters obtained were a learning rate of 0.00019685, dropout rate of 0.374138, the Adam optimizer,weight decay is 0.000118 and 512 hidden units in the fully connected layer. For the ResNet50 model, Optuna selected a learning rate of 0.00018047, dropout rate of 0.4386, the AdamW, weight decay is 1.1401322779456681e-05 and 1024 hidden units.

To address the class imbalance in the dataset, a weighted categorical cross-entropy loss function was employed. The training process incorporated several performance-enhancing techniques, including:

- **Early stopping:** To prevent overfitting and reduce unnecessary training time, the training was halted if the validation loss did not improve for a predefined number of epochs.
- **Learning rate scheduling:** The learning rate was adaptively adjusted during training to ensure stable convergence.
- **Stochastic Weight Averaging:** improves model generalization by averaging weights from multiple training epochs instead of using the final model weights. It finds wider optima that generalize better and is especially effective with high learning rates. SWA is simple to implement and requires no changes to model architecture.
- **Data augmentation:** Applied to minority classes during preprocessing to improve generalization (as detailed in the preprocessing section).

The models were trained on GPUs, and the training pipeline was designed to ensure reproducibility and efficient hyperparameter search using Optuna's Tree-structured Parzen Estimator (TPE) sampler.

F. Evaluation Metrics

Model evaluation was based on accuracy, precision, recall, and F1-score. Additionally, confusion matrices and training curves provided insight into performance.

IV. Results and Discussion

To evaluate our models, we trained both ResNet50 and DenseNet-121 on the processed Cat Breeds Dataset. The test results were evaluated using common metrics: accuracy, precision, recall, F1-score, and loss. Additionally, learning curves, confusion matrices, and ROC curves were analyzed.

A. Quantitative Comparison

Table I summarizes the performance of both models. ResNet50 slightly outperforms DenseNet-121 in accuracy, F1-score, and test loss.

TABLE I
Performance Comparison Between ResNet50 and DenseNet-121

Metric	ResNet50	DenseNet-121
Test Loss	1.2762	1.2910
Accuracy (%)	73	71
Precision	0.72	0.71
Recall	0.72	0.71
F1 Score	0.72	0.71

B. Loss and Accuracy Curves

The training and validation curves demonstrate convergence and effective learning. DenseNet-121 exhibits lower validation loss and better accuracy.

ResNet50:

DenseNet-121:

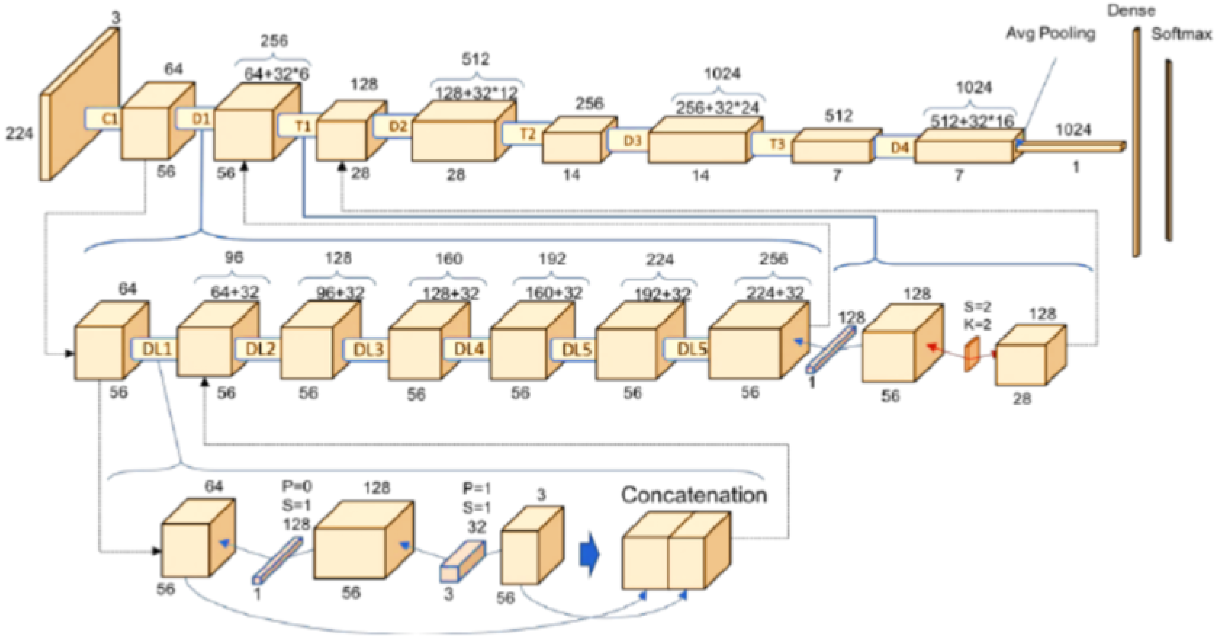
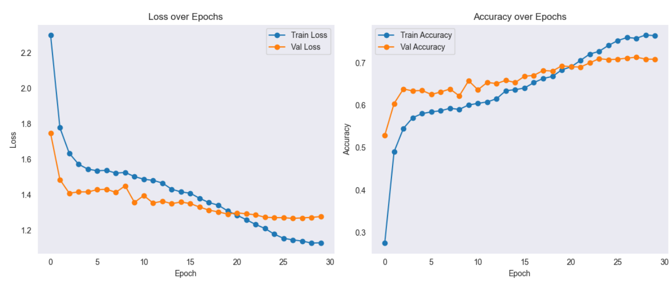
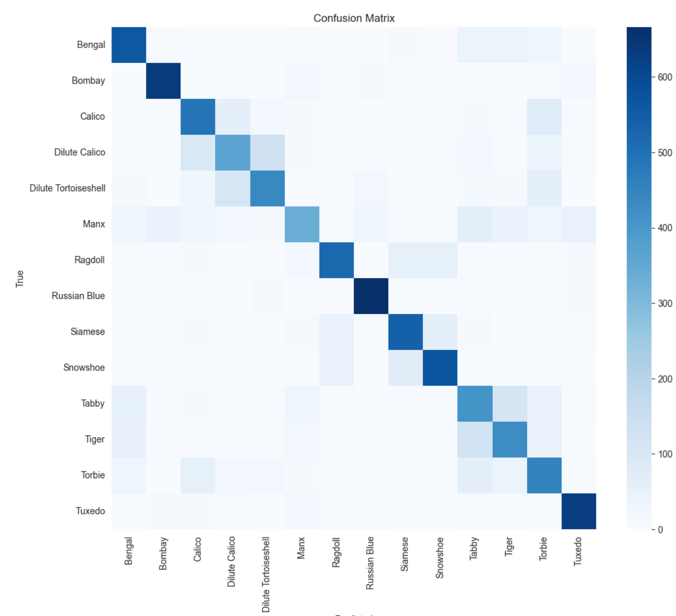


Fig. 4. Architecture of DenseNet-121



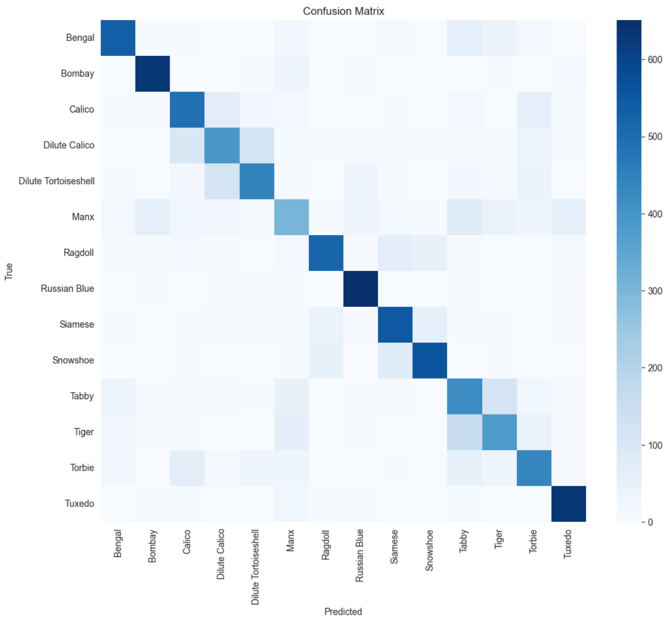


C. Confusion Matrices

The confusion matrices reveal that both models achieved high class-wise accuracy, with EfficientNet-B0 performing slightly better in minority classes.

ResNet50:

DenseNet-121:



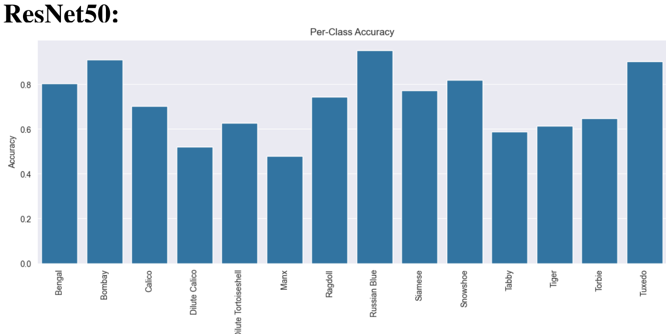
D. Per Class Accuracy

To calculate per-class accuracy, you use the formula:

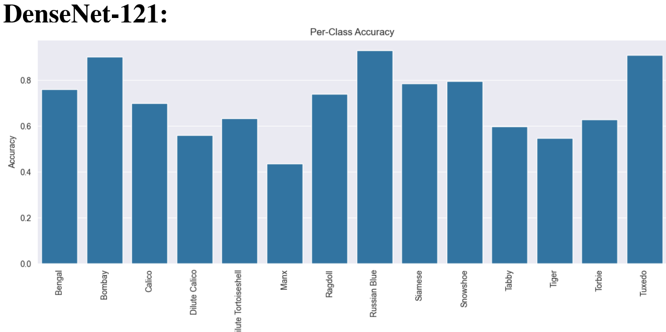
$$\text{Per-class Accuracy}i = \frac{\text{Correct Predictions for class } i}{\text{Total samples of class } i} = \text{Recall}_i$$

In the classification report you shared, the recall for each class is equivalent to per-class accuracy (because each class is treated as the "positive" class in a one-vs-rest evaluation).

ResNet50:



DenseNet-121:



E. Discussion

DenseNet-121's dense connectivity and efficient parameter usage contributed to improved feature reuse and generalization. The performance gains were further supported by:

- Optimal hyperparameter tuning using Optuna
- Weighted categorical cross-entropy loss
- Early stopping and learning rate scheduling
- Targeted data augmentation for underrepresented classes

These findings highlight the potential of deep learning models like DenseNet121 in enhancing automated cat breeds classification, particularly within the context of precision agriculture.

F. Limitations and Challenges

Despite the promising results, several limitations were identified during the course of this work:

- **Visual Similarity:** Many breeds are visually similar, causing confusion.
- **Dataset Coverage:** 14 classes, including some non-standard "breeds."
- **Data Diversity and Class Imbalance:** Mixed image quality, variety of settings, mislabels. Heavy skew toward a few dominant classes; many classes have very few images.

V. Conclusion and Future Work

This research introduced a deep learning-based system for cat breeds classifications using ResNet50 and DenseNet-121. Despite the similar performance of the two models, ResNet50 slightly outperformed DenseNet-121 for applications with limited resources, such as mobile farming tools. Even though the performance was strong, certain challenges such as class imbalance and the similarity of cat breeds.

As part of future work, we aim to extend the dataset to include more breeds, to enhance augmentation and domain adaptation strategies, achieve more accuracy, explore real-time deployment on smartphones, and to improve transparency and trust in the model.

References

[1] N. I. Progga, N. Rezoana, M. S. Hossain, R. U. Islam, and K. Andersson, "A CNN based model for venomous and non-venomous snake classification," in *Applied Intelligence and Informatics – First International Conference, AII 2021, Nottingham, UK, July 30–31, 2021, Proceedings*, ser. Communications in Computer and Information Science, vol. 1435. Springer, 2021, pp. 216–231.

[2] A. M. Durso, G. K. Moorthy, S. P. Mohanty, I. Bolon, M. Salathé, and R. Ruiz de Castañeda, "Supervised learning computer vision benchmark for snake species identification from photographs: Implications for herpetology and global health," *Frontiers in Artificial Intelligence*, vol. 4, p. 582110, Apr. 2021. [Online]. Available: https://doi.org/10.3389/frai.2021.582110

[3] L. Bloch and C. M. Friedrich, "Efficientnets and vision transformers for snake species identification using image and location information," in *Proceedings of the Working Notes of CLEF 2021 – Conference and Labs of the Evaluation Forum*, ser. CLEF '21. Bucharest, Romania: CEUR-WS.org, 2021, pp. 1477–1498, anthology ID: DBLP:conf/clef/BlochF21.

[4] K. Ahmed, M. A. Gad, and A. E. Aboutabl, "Snake species classification using deep learning techniques," *Multimedia Tools and Applications*, vol. 82, no. 16, p. 26021–26045, Sep. 2023. [Online]. Available: https://doi.org/10.1007/s11042-023-16773-0

[5] M. Rajabizadeh and M. Rezghi, "A comparative study on image-based snake identification using machine learning," *Scientific Reports*, vol. 11, no. 1, p. 19142, Sep. 2021. [Online]. Available: https://doi.org/10.1038/s41598-021-96031-1

[6] L. Kalinathan, P. Balasundaram, P. Ganesh, S. S. Bathala, and R. K. Mukesh, "Automatic snake classification using deep learning algorithm," in *Proceedings of the Working Notes of CLEF 2021 – Conference and Labs of the Evaluation Forum*, ser. CLEF '21. Bucharest, Romania: CEUR-WS.org, 2021, pp. 1587–1596, anthology ID: DBLP:conf/clef/KalinathanBGBM21.

[7] L. Picek, M. Hrúz, A. M. Durso, and I. Bolon, "Overview of snakeclef 2022: Automated snake species identification on a global scale," in *Proceedings of the Working Notes of CLEF 2022 – Conference and Labs of the Evaluation Forum*, ser. CLEF '22. Bologna, Italy: CEUR-WS.org, 2022, pp. 1957–1969.

[8] C. Zou, F. Xu, M. Wang, W. Li, and Y. Cheng, "Solutions for fine-grained and long-tailed snake species recognition in snakeclef 2022," *arXiv preprint*, vol. arXiv:2207.01216, Jul. 2022. [Online]. Available: https://arxiv.org/abs/2207.01216

[9] R. Chamidullin, M. Šulc, J. Matas, and L. Picek, "A deep learning method for visual recognition of snake species," in *Proceedings of the Working Notes of CLEF 2021 – Conference and Labs of the Evaluation Forum*, ser. CLEF '21. CEUR-WS.org, 2021, pp. 1512–1525.

[10] A. Patel, L. Cheung, N. Khatod, I. Matijošaitienė, A. Arteaga, and J. W. Gilkey Jr., "Revealing the unknown: Real-time recognition of galápagos snake species using deep learning," *Animals*, vol. 10, no. 5, p. 806, May 2020. [Online]. Available: https://doi.org/10.3390/ani10050806

[11] K. Desingu, M. Palaniappan, and J. Kumar, "Snake species classification using transfer learning technique," EasyChair Preprint, Tech. Rep. 7214, 2024, easyChair Preprint 7214.

[12] L. Bloch, A. Boketta, C. Keibel, E. Mense, A. Michailutschenko, O. Pelka, J. Rückert, L. Willemeit, and C. M. Friedrich, "Combination of image and location information for snake species identification using object detection and efficientnets," in *Working Notes of CLEF 2020 – Conference and Labs of the Evaluation Forum*, ser. CLEF '20. Thessaloniki, Greece: CEUR-WS.org, Sep. 2020, p. –, fHDO Biomedical Computer Science Group (BCSG).

[13] M. Palaniappan, K. Desingu, H. Bharathi, E. A. Chodisetty, and A. Bhaskar, "Deep learning and gradient boosting ensembles for classification of snake species," in *Proceedings of the Working Notes of CLEF 2022 – Conference and Labs of the Evaluation Forum*, ser. CLEF '22. Bologna, Italy: CEUR-WS.org, 2022, pp. 2175–2188.

[14] L. Bloch, J. Böckmann, B. Bracke, and C. M. Friedrich, "Combination of object detection, geospatial data, and feature concatenation for snake species identification," in *Proceedings of the Working Notes of CLEF 2022 – Conference and Labs of the Evaluation Forum*, ser. CLEF '22. Bologna, Italy: CEUR-WS.org, 2022, pp. 1982–2013.