Multi-Class Animal Image Classification Using Transfer Learning with VGG16 and InceptionV3

Dr. Manjunath H, Shreya Hegde, Shreya R, Varshini D, Yeshwanth S
Department of Computer Science and Engineering Bangalore Institute of Technology K.R. Road, V.V. Pura, Bengaluru - 560004, India.

Article Info

Keywords:

Transfer learning
VGG16
InceptionV3,
image classification,
deep learning,
performance metrics
confusion matrix

ABSTRACT

There is a growing need for automated classification of animal species, particularly in wildlife conservation, ecological research, environmental monitoring, and precision agriculture. Manual classification methods are time-consuming, labor-intensive, and can be prone to human error when dealing with large datasets; therefore, automated intelligent systems provide a significant advantage. This work presents a framework to address these challenges using a multi-class image classification approach based on transfer learning and deep learning methodologies. Two widely accepted convolutional neural network (CNN) architectures, VGG16 and InceptionV3, are employed to classify animal images into six categories: butterflies, chickens, elephants, horses, spiders, and squirrels. Pre-processing steps such as resizing, normalization, and data augmentation are systematically applied to enhance dataset quality and improve the generalization capability of the models. Both models are finetuned with customized classification layers to efficiently extract meaningful features while reducing computational overhead. Model performance is evaluated using accuracy metrics, loss analysis, confusion matrices, and detailed classification reports to assess prediction reliability. Experimental findings indicate that although both VGG16 and InceptionV3 demonstrate strong classification performance, InceptionV3 achieves superior accuracy, faster convergence, and greater stability due to its factorized convolutional architecture and multi-branch design.

Corresponding Author:

Yeshwanth S
Department of Computer Science and Engineering, Bangalore Institute of Technology
K. R. Road, V. V. Pura, Bengaluru- 560 004
Email: yeshwanthgowda6549@gmail.com

1. INTRODUCTION

Correct species identification still holds prominence in biodiversity, ecological monitoring, precision agriculture, and automated management of wildlife species. Manual species labeling of large image collections of wild animals, especially through camera traps, is still a cumbersome process, thereby justifying the use of automated species classification systems based on deep learning techniques. Early works proved the applicability of the concept of transfer learning using advanced convolutional networks such as Xception and VGG16 for the automated species classification of visually similar endangered species and primates, without requiring large amounts of data for the project's successful completion [1],[2]. Further developments in dog breed classification and multi-species classification brought about the significance of hybrid approaches using deep learning techniques for enhanced precision, generalization, and applicability [3],[4]. Animal species identity detected through aerial maps and drones was found extremely reliable through application of multi-class recognition models using deep learning techniques, ensuring successful species recognition in adverse environmental conditions [5]. Multi-class object recognition, species monitoring tasks, and improved prediction models using advanced CNN techniques and other species recognition systems further brought about the surplus applicability of techniques using CNN models in ecological systems [6],[7],[10]. Application of concepts using domain-aware preprocessing and domaGoogLeNet, DenseNet, and ResNet, supported by preprocessing models based on domain intelligence [8],[9],[11]–[13]. Further advancements were made using difficult-to-recognize species like aquatic, visually resembling parrot species, and small data sets, thereby validating the applicability of CNN technique concepts using challenging recognition methods [14], [15]. Further recent advancements brought about enhanced precision using species- ResNet, enhanced recognition systems using wild species, and automated systems for ecological environmental monitoring solutions through species recognition using interfaced systems [16]–[20]. Further developments were sought using species monitoring techniques like behavioral monitoring through wearable techniques, semisupervised models, species recognition through tactile techniques, difficult species location algorithms, and smart warning systems based on intelligent algorithms, thereby validating further advancements using ecological deep learning concepts and systems for species monitoring [21]–[25].

- To achieve better multi-class image classification performance by applying transfer learning using VGG16 and InceptionV3 models.
- To design an optimized deep learning pipeline to further improve feature extraction for better overall model performance in image classification tasks.
- To implement and evaluate a transfer learning–based approach using VGG16 and InceptionV3 to achieve the best results in the classification on a given dataset.

The literature survey in Section II outlines the motivation behind selecting transfer learning for multi-class animal classification. In Sections III, it is described the proposed methodology, dataset preparation, architectural design, algorithms, and mathematical formulations. Section IV reports the training-related experimental results of VGG16 and InceptionV3, followed by Section V, which concludes the study and identifies possible avenues for further research.

2. METHOD

A. Dataset Collection and Preprocessing

A curated dataset containing 17,803 pictures were gathered from publicly accessible sources such as Kaggle and additional online repositories. The images belong to six classes: butterfly, chicken, elephant, horse, spider, and squirrel. Each sample was manually inspected to remove corrupted or low-quality images. The dataset was divided into training (70%), validation (15%), and testing (15%) sets to maintain consistency for evaluation.

Preprocessing ensures that all images follow a uniform structure before being used by VGG16 and InceptionV3.Image Loading: Images are loaded from class-specific directories using TensorFlow's ImageDataGenerator. Resizing: VGG16 inputs are resized to 224 224 pixels, while InceptionV3 inputs are resized to 299 299 pixels to match their architecture requirements. Array Conversion: Each image is converted to a numerical array using img_to_array(). Batch Dimension: A batch axis is added using expand dims() since CNNs operate on batches. Normalization: The preprocess_input() functions apply ImageNet-based scaling and channel normalization. Data Aug- mentation: Random transformations such as rotation, zoom, flipping, shear, and brightness adjustments are applied during training to increase variability and reduce overfitting. These steps standardize the dataset and support stable convergence for both VGG16 and InceptionV3 during training.

B. Model Architecture

1) VGG16 Architecture: Figure 1 illustrates the architecture of the VGG16-based CNN used for multi-class image classification. The model processes a 224 224 3 RGB image via a sequence of convolutional and pooling layers, followed by layers that are fully connected and produce class probabilities via an activation function for SoftMax.

   The feature extraction backbone consists of five convolutional blocks that progressively learn hierarchical representations:

   - Block 1: Two 3 3 convolutional layers with 64 filters, then succeeded by 2x2 max pooling to capture basic edges and textures.
   - Block 2: Two convolutional layers with 128 filters, followed by max pooling for more structured feature extraction.
   - Block 3: Three layers of convolution with 256 filters, then pooling, enabling mid-level feature learning.
   - Block 4: Three layers of convolution with 512 filters and pooling, capturing high-level semantic features.
   - Block 5: Three layers of convolution with with 512 filters, followed by max pooling to produce compact feature maps.

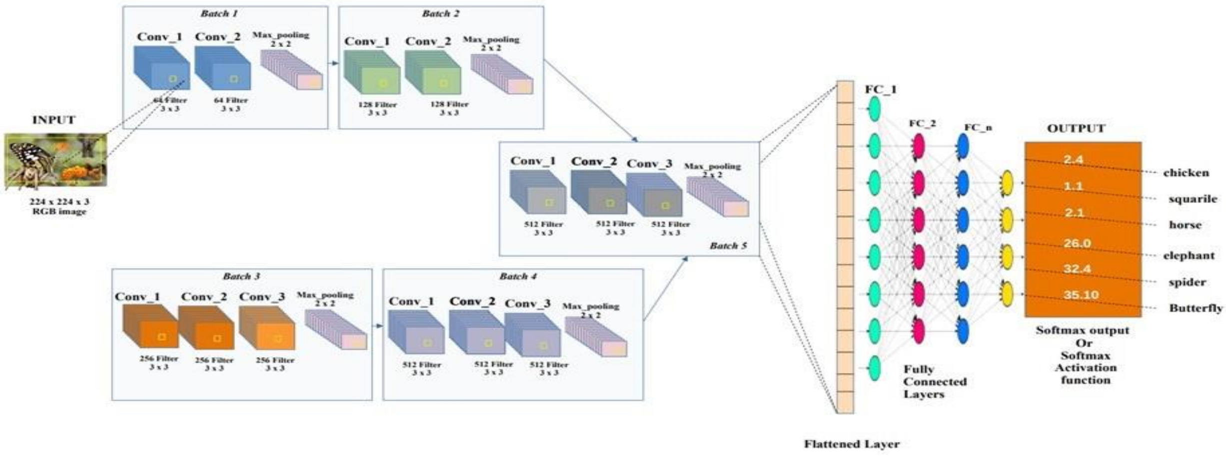


Figure 1.VGG16 architecture for the purpose of extracting features and classification.

2) InceptionV3 Architecture: Figure 2 illustrates the InceptionV3-based architecture used in this work. The network accepts a 299 × 299 × 3 RGB input image and processes it through an initial convolutional stem, followed by a sequence of stacked Inception modules, a Global Average Pooling layer, and fully connected layers.

The Inception modules are organized into the following groups:

- Mixed 5a–5c (Early Modules): Capture medium-level visual patterns such as textures and shapes using parallel convolutional branches.
- Mixed 6a–6f (Intermediate Modules): Learn more abstract and discriminative features, improving robustness to image scale and rotation variations.
- Mixed 7a–7b (Late Modules): Encode high-level semantic information that is essential for final classification.

A Global Average Pooling layer converts the 2D feature maps into a compact 1D vector representation, which is then passed to a fully connected layer followed by a final SoftMax layer to generate probability scores for the six target classes.

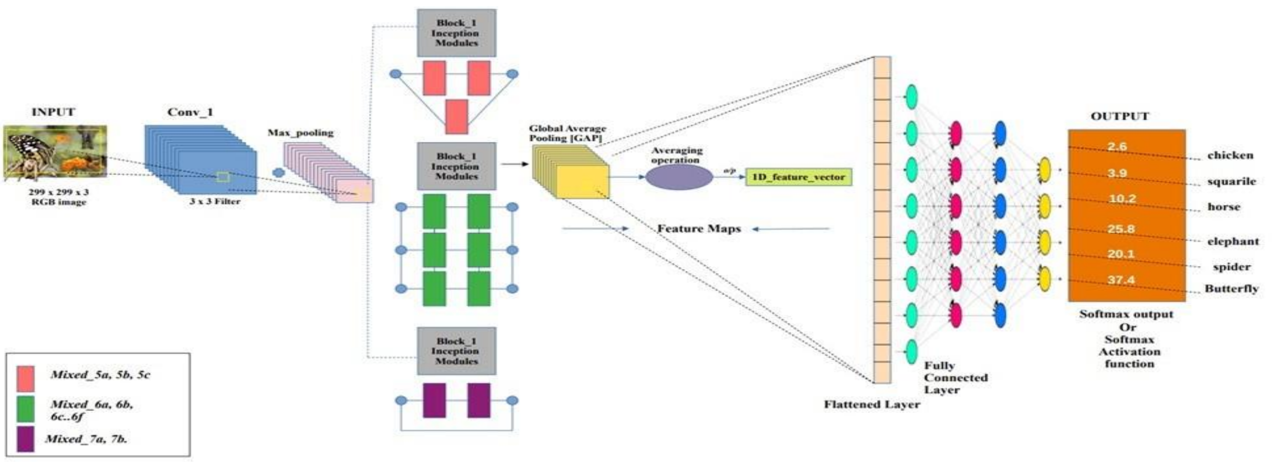


Fig. 2: InceptionV3 architecture with multi-scale Inception modules.

C. Mathematical Equations

1) VGG16

Input (X): The resized and normalized image is passed into the VGG16 model.

$$F = \mathrm{VGG16}(X)$$

Flatten Feature map: Flatten the feature map into a 1D vector.

$$F_{\mathrm{fa}} = \mathrm{Flatten}(F)$$

Dense Layer and Activation: passes the flattened feature vector f through a series of Dense Layer.

First Dense Layer:

$$Z = W \cdot F_{\mathrm{fa}} + b$$

ReLU Activation:

$$A = \mathrm{ReLU}(Z)$$

Cross Entropy Loss Function : During the training, the model minimizes the cross-entropy loss

$$L = -\sum_{c=1}^{K} y_c \cdot \log\big(P(c)\big)$$

Backpropagation and optimization: using backpropagation, the weights are updated to minimizes the loss

$$\mathrm{W} = \mathrm{W} - \eta \cdot \frac{\partial L}{\partial \mathrm{W}}$$

Final output layer (Softmax): The Logits are converted into probabilities using Softmax functions.

$$P(c) = \frac{e^{Z_c}}{\sum_{k=1}^{K} e^{Z_k}}$$

Final Prediction: For a new Image, the predicted class given by

$$\hat{c} = \arg\max\big(P(c)\big)$$

Final Formula: Bringing it all together, the final formula for the VGG16 multi-classification process is:

$$\hat{c} = \arg\max\left(\frac{e^{W\cdot \mathrm{Flatten}\left(\mathrm{VGG16}(X)\right)+b_c}}{\sum_{k=1}^{K} e^{W\cdot \mathrm{Flatten}\left(\mathrm{VGG16}(X)\right)+b_k}}\right)$$

- X is the input image
- W are weights of dense layer
- VGG16(X) represents feature extraction
- The ReLU activation introduces non-linearity
- Softmax normalizes the logits the probabilities.
- Argmax selects the class with the highest probability

2) INCEPTIONV3

Each input image X is resized, normalized, and reshaped to match the InceptionV3 input dimensions (299x299x3)

$$X' = \frac{X}{255}, \quad X' \in R^{299\times299\times3}$$

The pre-processed image X' passes through the InceptionV3 layers, producing a feature map

$$F = InceptionV3(X')$$

The feature map F is reduced using Global Average Pooling to a vector F_{GAP}, condensing each feature map's information into a single value

$$F_{\text{GAP}} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_{i,j,k}$$

The pooled features F_{GAP} are passed through a fully connected (Dense) layer to produce the final class logits Z

$$Z = W \cdot F_{\text{GAP}} + b$$

To convert the logits into probabilities, we apply the softmax function:

$$P(y_i \mid X') = \frac{\exp(Z_i)}{\sum_{j=1}^{N} i \exp(Z_j)}$$

The model's performance is measured using categorical cross-entropy, comparing the predicted $P(y_i \mid X')$

$$\mathcal{L} = -\sum_{i=1}^{N} y_i \log\left(P(y_i \mid X')\right)$$

The weights W are updated using the Adam optimizer, which combines momentum and adaptive learning rates:

$$W_{t+1} = W_t - \alpha \cdot \frac{\widehat{m_t}}{\sqrt{\widehat{v_t}} + \epsilon}$$

For a new test image X_{test}, the model predicts the class probabilities:

$$P(y_i \mid X_{\text{test}}) = \text{softmax}(Z_{\text{test}})$$

Summarizing all steps, the entire process from input to prediction can be written as

$$P(y_i \mid X) = \text{softmax}\left(W \cdot \text{GAP}\left(\text{InceptionV3}\left(\frac{X}{255}\right)\right) + b\right)$$

3. RESULTS AND DISCUSSION

3.1. Training Curves

This subsection presents a thorough examination of the learning behavior of both VGG16 and InceptionV3 during training. To maintain clarity, each explanation is immediately followed by its corresponding accuracy or loss plot.

3.1.1. VGG16 – Train and Validation Accuracy: The acc graph for VGG16 indicates that our models tend to learn very well. First, the training accuracy begins lower due to model having yet learned basic features from the data. But for each epoch, the average training/validation accuracy increase over epochs means that the model is “learning” to identify animals from different classes more effectively across time. The little backbone is normal to see pull in validation accuracy due to some discrepancy between the training data and the unseen validation data. Nevertheless, despite these oscillations we can observe the general tendency-upward trend which indicates that learned features are semantical and after many samples the network becomes better able to distinguish classes rather than making random guesses

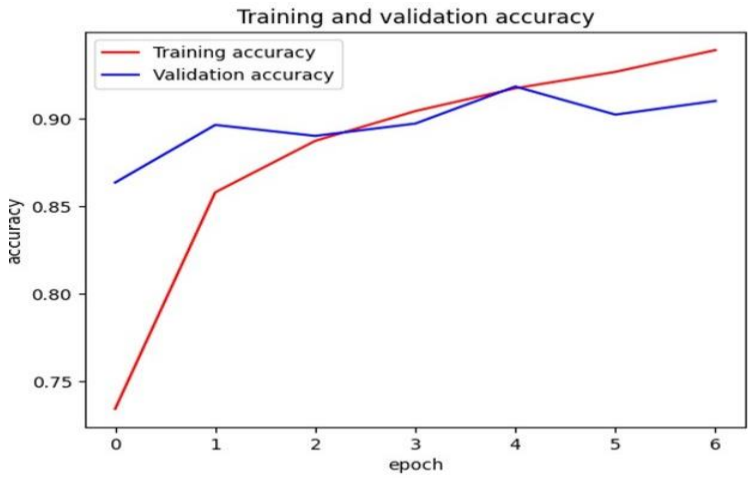


Fig. 3: VGG16: Training and validation accuracy curve.

3.1.2. VGG16 – Training and Validation Loss:The loss curves will also help to understand whether the model is learning efficiently or not. Training loss curves for VGG16: It is clear to see that the losses reduces in a smooth manner with the increasing epochs which means that the model still seems to be improving its predictions. The validation loss decreases as well for some number of layers and then stays the same. This convergence indicates that the model is close to its maximum learning capacity for the dataset. Beyond this point, overtraining becomes a problem, where the model learns from noise in training data. The fact that the curves are close to each other and they decrease smoothly suggests that training is reasonably well stabilized, and the model parameters selected were good.

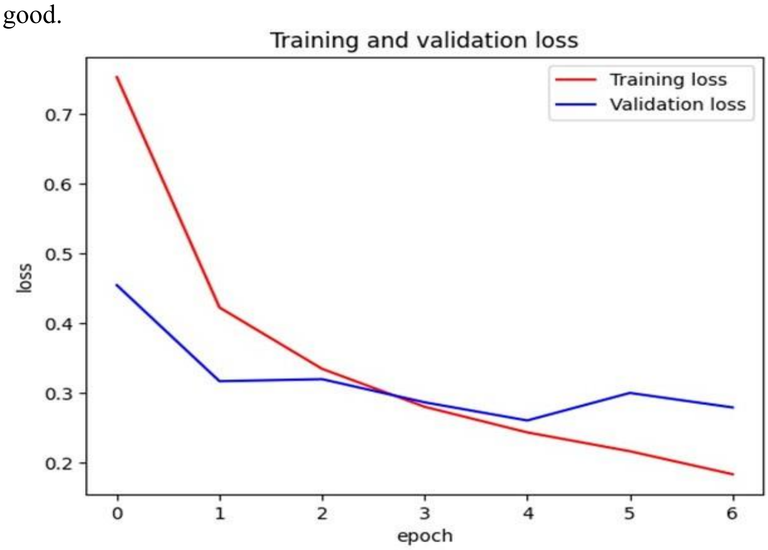


Fig. 4: VGG16: Training and validation loss curve.

3.1.3.InceptionV3 – Training and Validation Accuracy: InceptionV3 has a significantly better convergence than VGG16. The trend continues along the training set where we obtain very high accuracy rates at a very early stage of the training. This result demonstrates also the effectiveness of InceptionV3 architecture, capable to capture complex image patterns. Probably, the most important is that, as shown below, the validation accuracy follows very closely the training one without a big gap between them. This indicates a great generalization ability i.e that the model is doing good not only on images used for training but also on unseen validation images. The consistently-high accuracy values indicate that InceptionV3 is capable of treating multi-class classification effectively and can generate trustable predictions.

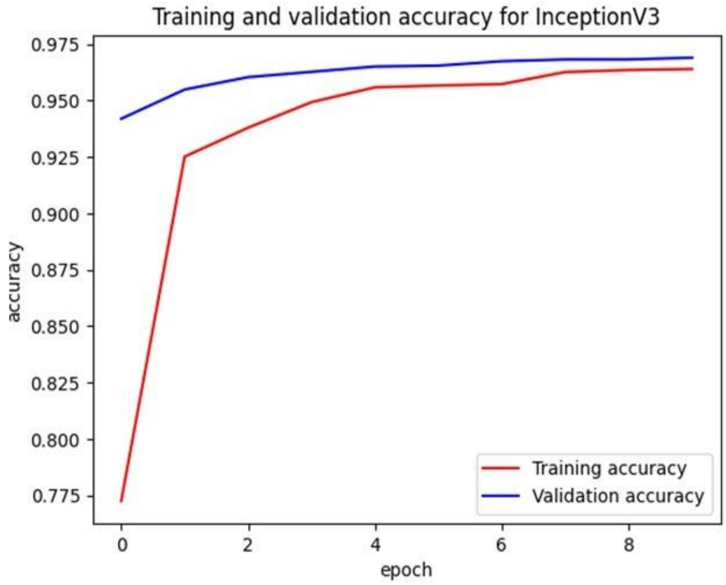
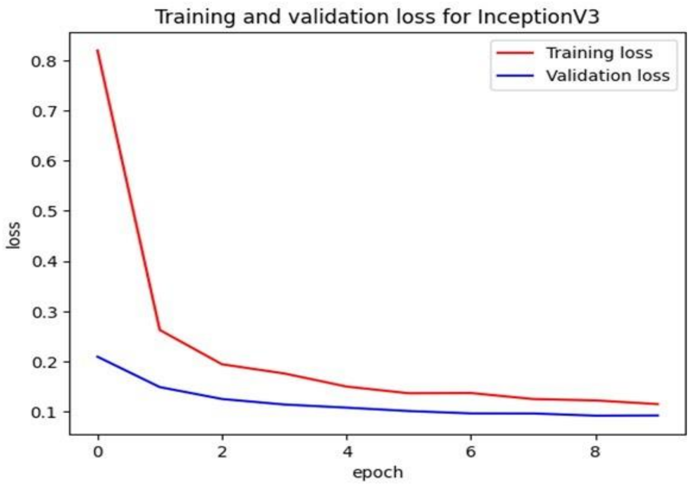


Fig. 5: InceptionV3: Training and validation accuracy curve.

3.1.4. InceptionV3 – Training and Validation Loss: For InceptionV3, the loss curve plummets very fast as observed in some initial epochs, demonstrating fast learning rate and optimum weight adjustments. The validation loss is staying so close to the training loss for all of the training processes that this suggests very low levels of overfitting. This stability demonstrates that the model generalizes well by learning patterns and not relying much on a specific training example. The final loss values are consistently small, which suggests that the model converges at a confident learning state. In the larger context, the behaviors of these curves indicate that InceptionV3 learns faster and also keeps good stability as well as balance between learning and generalization.



2) Fig. 6: InceptionV3: Training and validation loss curve

3.2. Confusion Matrices

Confusion matrices provide fine-grained insights into model performance by illustrating how often predictions match ground-truth labels at a per-class level. They also help identify patterns of misclassification, especially among visually similar animal categories.

3.2.1. VGG16 Confusion Matrix: The confusion matrix of VGG16 indicates that the model works well for six animal classes, with a majority of predictions along the diagonal line (i.e. correct classification). High diagonal values show that the model has learnt a good set of visual features for both butterflies, elephants, spiders and squirrels — resulting in strong recognition performance on these object categories. But there are a few of apparently misclassifications. Horses and elephants are not infrequently mistaken for other classes. This can be the result of similarity in visual patterns (e.g., textures) or because of noise contained in some images. Also, VGG16 has an older architectural design and most of the time require deeper stacks of convolutions that might not be as sensitive to fine-grained differences in some other more modern architectures. Despite these issues, we notice a well equilibrated distribution over the predictions and no class is strongly misclassified. This suggests that VGG16 can serve as a reliable and consistent approach for multi-class AIC, yet its own performance also demonstrates some room for improvement, especially against visually similar or complex samples.

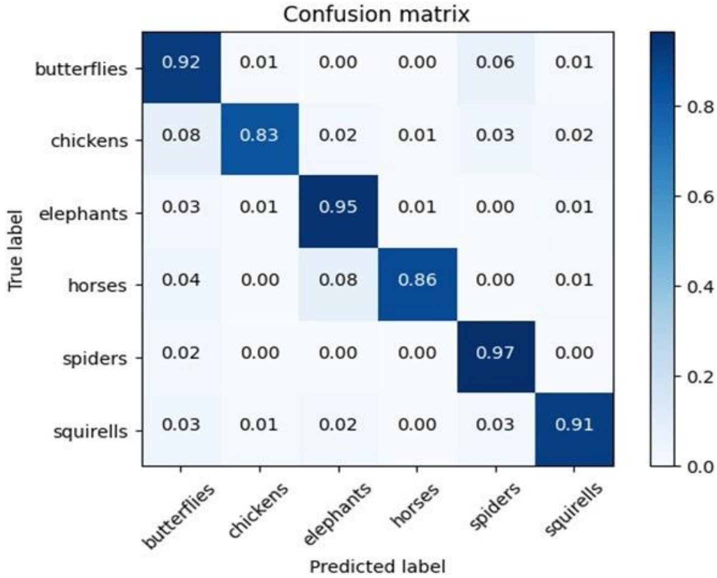


Fig. 7: VGG16: Normalized confusion matrix.

3.2.2. InceptionV3 Confusion Matrix: The confusion matrix for InceptionV3 shows a very strong classification capability, with nearly all of the values on diagonal. That is, for the 399,787 test samples in total spanning all six categories most are prediction correctly with very little confusion. The upper field of the table appears to be nearly all close to 1 in a sense on classes butterfly, chicken, elephant, horse, spider and squirrel which again indicates the excellent recognition capability. This betterment over the VGG16 has much to do with InceptionV3's architectural capabilities. The factorized convolutions, parallel convolutional branch and multiscale features extraction in the network structure allows to better capture fine textures, shapes and context information. Therefore, InceptionV3 can recognize animals better with similar visual properties or in complex environments. These off-diagonal low values mean here few mistakes, assuring that the model is well generalizing and has adequate stability during predictions. This makes InceptionV3 particularly well-suited for real-life animal classification applications, when precision and performance are paramount.

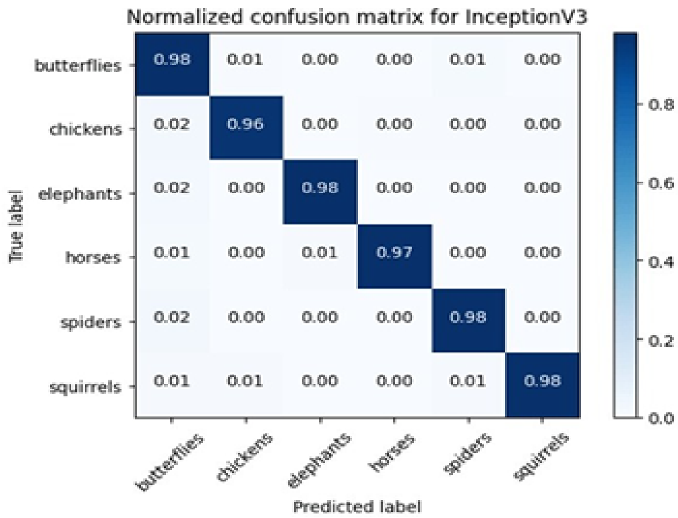


Fig. 8: InceptionV3: Normalized confusion matrix

3.3 Classification Metrics

3.3.1. VGG16 Classification Metrics: Table I presents the performance of the VGG16 model on all six animal classes. The performance metrics reveal a robust performance of the model with a weighted F1 score of 0.91, balancing its precision and recall rates. The accuracy of classes including butterflies, horses, and spiders is remarkably high, signifying excellent feature extraction capabilities of the model. However, a slight decrease in the recall rate of classes like chickens and elephants implies some overlap between classes due to image similarity. However, a robust performance of the VGG16 model has been noticed on 1,845 test examples of animal images with minimal errors and optimal accuracy for classification.

TABLE I: Classification Metrics (VGG16)

Class	Precision	Recall	F1-score	Support
butterflies	0.98	0.89	0.93	371
chickens	0.95	0.83	0.89	203
elephants	0.73	0.95	0.83	152
horses	0.99	0.86	0.92	472
spiders	0.91	0.97	0.94	403
squirrels	0.93	0.91	0.92	244
micro avg	0.93	0.90	0.91	1845
macro avg	0.92	0.90	0.90	1845
weighted avg	0.94	0.90	0.91	1845
samples avg	0.90	0.90	0.90	1845

3.3.2. InceptionV3 Classification Metrics: Table II represents the classification output of the InceptionV3 model, which ensures a considerable boost in the classification accuracy in comparison to VGG16. The model achieves 98% overall accuracy along with higher precision and recall points for all classes, resulting in a weighted F1-score of 0.98. An overall correctness in classification with utmost consistency in all classes signifies that the model predicts correctly and performs consistently on all classes. Moreover, the fact that the macro average and the weighted average are quite close ensures that the InceptionV3 model performs uniformly on all classes without any biasness.

TABLE II: Classification Metrics (InceptionV3)

Class	Precision	Recall	F1-score	Support
butterflies	0.98	0.97	0.97	371
chickens	0.97	0.97	0.97	203
elephants	0.99	0.98	0.98	152
horses	0.99	0.98	0.99	472
spiders	0.97	0.98	0.98	403
squirrels	0.97	0.99	0.98	244
accuracy		0.98		1845
macro avg	0.98	0.98	0.98	1845
weighted avg	0.98	0.98	0.98	1845
samples avg	0.98	0.97	0.98	1845

4. CONCLUSION

In this paper, two of the most common deep learning architectures, VGG16 and InceptionV3, will be compared in their performance for multi-class animal image classification. VGG16 is a deep convolutional neural network with an elegant stacked convolution and pooling structure that enables the successful extraction of strong hierarchical features from images. However, being of a large number of parameters means that more computation must be required and it will converge slightly slower. In contrast, InceptionV3 is an advanced architecture designed to achieve efficiency and gain accuracy with factorized convolutions, parallel processing branches, and multiscale feature extraction. This makes it capture finer visual details while maintaining faster training and better stability for convergence. When comparing both models, VGG16 showed a very good performance and proved that transfer learning can be super effective even when the available dataset is not enormous. It was able to achieve high classification accuracy but showed a relatively higher risk of overfitting and slower progress in learning. InceptionV3 performed better overall: faster convergence, it produced high validation accuracy, and minimal overfitting due to its architectural enhancements. Experimentally, InceptionV3 yielded a better accuracy with a stronger generalization capability than VGG16, thus supporting this method in dealing with complex and diverse image datasets. In general, VGG16 serves as a strong and stable baseline, while InceptionV3 improves the performance, stability, and confidence of the predictions. Further work may be related to investigating the performance of more recent architectures like EfficientNet or Vision Transformers, while further exploring lightweight deployment techniques so that these models could be applied effectively in real-time wildlife monitoring scenarios and other real-world applications.

ACKNOWLEDGMENTS

The authors express sincere gratitude to the Department of Computer Science and Engineering at the Bangalore Institute of Technology for providing infrastructure and academic support. The authors thank our guide, Dr Manjunath H, for guidance and feedback. We also acknowledge the contributions of peers who assisted with data set curation and experimentation and our families for their support during this work

FUNDING INFORMATION

This work was carried out as part of an academic research project under the Department of Computer Science and Engineering, without any external funding support.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Dr. Manjunath H	✓	✓			✓		✓			✓		✓	✓	
Shreya Hegde	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓		✓	
Shreya R		✓	✓	✓		✓		✓	✓	✓	✓			
Varshini D		✓		✓		✓	✓	✓	✓	✓	✓			
Yeshwanth S	✓		✓	✓	✓	✓	✓		✓					

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY

The dataset supporting the findings of this study consists of openly available animal images collected from public sources. Due to compilation licensing constraints, the final combined dataset is not publicly shared. However, the dataset and preprocessing files can be provided by the corresponding author upon reasonable request.

REFERENCES

[1] "Classification of Endangered Species of Wild Cats through Xception Transfer Learning Model," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10201732

[2] "Primate Recognition System Design Based on Deep Learning Model VGG16," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/9886310

[3] B. Valarmathi *et al.*, "Hybrid Deep Learning Algorithms for Dog Breed Identification — A Comparative Analysis," *IEEE Access*, vol. 11, pp. 77228–77239, 2023, doi: 10.1109/ACCESS.2023.3297440.

[4] "An Efficient Framework for Animal Breeds Classification Using Semi-Supervised Learning and Multi-Part CNN," *IEEE Journals & Magazine*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8877750

[5] "Deep Neural Networks Based Multiclass Animal Detection and Classification in Drone Imagery," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10323685

[6] "A Deep Learning Framework of Convolutional Neural Network Utilization for Multi-Class Object Recognition," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8492777

[7] "Classification of Animals with Different Deep Learning Models," *Journal of New Results in Science*, Accessed: Dec. 11, 2024. [Online]. Available: https://dergipark.org.tr/en/pub/jnrs/issue/36616/387258

[8] "Improving the Accuracy of Animal Species Classification in Camera Trap Images Using Transfer Learning," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10467777

[9] "Multispecies Bioacoustic Classification Using Transfer Learning of Deep CNNs with Pseudo-Labeling," *Applied Acoustics*, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0003682X20304795

[10] "An Enhanced Animal Species Classification and Prediction Engine Using CNN," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10212299

[11] P. A. Mary G *et al.*, "Wild Animal Detection System," in *2023 2nd Int. Conf. on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, pp. 1–6, 2023, doi: 10.1109/ICAECA56562.2023.10199701.

[12] "Breakthrough Conventional Based Approach for Dog Breed Classification Using CNN with Transfer Learning," *IEEE Conference Publication*. [Online]. Available: https://ieeexplore.ieee.org/document/8929955

[13] "A Small Sample Image Recognition Method Based on ResNet and Transfer Learning," *IEEE Conference Publication*. [Online]. Available: https://ieeexplore.ieee.org/document/9178671

[14] W. Jang and E. C. Lee, "Multi-Class Parrot Image Classification Including Subspecies with Similar Appearance," *Biology*, vol. 10, no. 11, 2021, doi: 10.3390/biology10111140.
[15] "Aquatic Animal Image Classification Technology Based on Transfer Learning and Data Augmentation," *Journal of Coastal Research*. [Online]. Available: https://meridian.allenpress.com/jcr/article-abstract/105/SI/129/450191/

[16] "Animal Species Classification Using Fine-Tuned ResNet152V2: A Deep Learning Approach for Biodiversity Conservation," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/11011471

[17] "High-Accuracy Animal Species Classification Using Transfer Learning with ResNet50 on Wildlife Camera Trap Images for Enhanced Conservation Efforts," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10960173

[18] "Recognizing Wild Animals from Camera Trap Images Using Deep Learning," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: httpsieeexplore.ieee.org/document/9774826

[19] "Wildlife Species Recognition Using Deep Learning," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10394465

[20] "Improving the Accuracy of Animal Species Classification in Camera Trap Images Using Transfer Learning," *IEEE Conference Publication*, Accessed: Dec. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10467777

[21] A. L. Hernández, M. Á. F. Hernández, R. García, and C. J. Pérez,"Activity Detection for the Wellbeing of Dogs Using Wearable Sensors Based on Deep Learning," IEEE Sensors Journal, vol. 22, no. 14, pp. 14636–14645, 2022,doi: 10.1109/JSEN.2022.3181234.

[22] S. Alirezazadeh, M. H. Moradi, and S. M. R. Soroushmehr, "An Efficient Framework for Animal Breeds Classification Using Semi-Supervised Learning and Multi-Part Convolutional Neural Network (MP-CNN), "IEEE Access, vol. 8, pp. 155443–155456, 2020, doi: 10.1109/ACCESS.2020.3019174.

[23] Z. Abderrahmane, G. Ganesh, A. Crosnier, and A. Cherubini, "A Deep Learning Framework for Tactile Recognition of Known as Well as Novel Objects," IEEE Transactions on Industrial Informatics, vol. 16, no. 1, pp. 423–432, Jan. 2020, doi:10.1109/TII.2019.2898264.

[24] H. Teigen, A. Ahmed, A. S. Imran, M. Ullah, R. M. A. Azad, and A. Soylu,"From Blurs to Birds: Localization and Classification of Hard-to-See Bird Species in Norwegian Wilderness Camera Trap Images," IEEE Access, vol. 13, pp. 169534–169553, 2025, doi: 10.1109/ACCESS.2025.3613068.

[25] B. Natarajan, R. Elakkiya, R. Bhuvaneswari, K. Saleem, D. Chaudhary, and S. H. Samsudeen, "Creating Alert Messages Based on Wild Animal Activity Detection Using Hybrid Deep Neural Networks," IEEE Access, vol. 11, pp. 67308–67321, 2023, doi: 10.1109/ACCESS.2023.3289586.

BIOGRAPHIES OF AUTHORS

Dr. Manjunath H is an Assistant Professor in the Department of Computer Science and Engineering. He holds a B.E., M.Tech., and Ph.D., with specialization in Database Management Systems, Data Structures using C, Algorithms, and Software Engineering and Management. His teaching expertise covers Database Management Systems, Programming in C and Data Structures, Object-Oriented Concepts, Programming in C for Problem Solving, Analysis and Design of Algorithms, Advanced Algorithms, Software Engineering, Entrepreneurship, and User Interface Design. His research interests primarily focus on Machine Learning and Artificial Intelligence, with emphasis on developing intelligent computational approaches to address real-world challenges. He has published three research papers in recognized journals and conferences, contributing to advancements in computing and intelligent systems. He remains actively engaged in guiding students and supporting innovative research and academic development. He can be contacted at email: manjunathh@bit-bangalore.edu.in

Shreya Hegde is an undergraduate student in the Department of Computer Science and Engineering with a strong academic foundation in artificial intelligence, machine learning, and data-driven computing. Her academic training emphasizes deep learning, computer vision, and intelligent systems development. She has been involved in academic and project-based research works focusing on AI-enabled applications, multi-class image classification, and intelligent automation systems, contributing to design, implementation, and evaluation activities. She continues to pursue advancements in AI and computer vision to contribute impactful research to the scientific community. She can be contacted at email: shreya4hegde@gmail.com

Shreya R . Shreya R is an undergraduate student in the Department of Computer Science and Engineering with a strong academic background in artificial intelligence, machine learning, and data-driven computing. Her academic training focuses on deep learning, computer vision, and intelligent system development, with emphasis on efficient and scalable computational models. Her research interests include transfer learning, neural network optimization, model generalization, and the application of artificial intelligence to real-world problem solving. She has developed a sound understanding of machine learning workflows, data preprocessing techniques, and performance evaluation of deep learning models, along with interests in intelligent automation and pattern recognition She can be contacted at email: shreyar24904@gmail.com

Varshini D is an undergraduate student in the Department of Computer Science and Engineering with a strong academic foundation in artificial intelligence and machine learning. Her academic focus includes deep learning and computer vision, and her research interests involve transfer learning, model optimization, and real-world AI applications. She has participated in academic and project-based research related to AI-enabled systems and multi-class image classification. In this paper, she contributed to defining the research problem, conducting the literature survey, preparing the dataset, and selecting transfer learning models such as VGG16 and InceptionV3. She was involved in model training, experimental analysis, and comparative evaluation. She also contributed to drafting and organizing the manuscript, ensuring technical clarity and accuracy. She continues to pursue research in artificial intelligence and computer vision to contribute meaningful work to the research community. She can be contacted at email: varshini4123@gmail.com

Yeshwanth S is an undergraduate student in the Department of Computer Science and Engineering with a strong academic foundation in artificial intelligence, machine learning, and data-driven computing. His academic training focuses on deep learning, computer vision, and the development of intelligent systems. He has a keen research interest in transfer learning, neural network performance optimization, and improving model generalization for practical applications. Yeshwanth has been actively involved in academic and project-based research, gaining hands-on experience in designing and implementing AI-enabled solutions, including multi-class image classification and other real-world problem-solving applications. His work reflects a clear balance between theoretical knowledge and practical implementation, demonstrating his ability to apply advanced AI techniques to address complex computational challenges effectively.She can be contacted at email: yeshwanthgowda6549@gmail.com