# Hybrid Approach For Animal Breed Classification Using EfficientNet: A Deep Learning Model For Multi-Class Image Recognition

1st K Venkatesan
*Dept. of Computer Science and Engg.*
*Amrita School of Computing*
*Amrita Vishwa Vidyapeetham*
Chennai,India

2nd K Deepak
*Dept. of Computer Science and Engg.*
*Amrita School of Computing*
*Amrita Vishwa Vidyapeetham*
Chennai,India

3rd P Sai Manoj Kumar
*Department of CSE(AI)*
*Amrita School of Computing*
*Amrita Vishwa Vidyapeetham*
Chennai,India

4th Pothala Mokshajna Venkata Krushna
*Department of CSE(AI)*
*Amrita School of Computing*
*Amrita Vishwa Vidyapeetham*
Chennai,India

***Abstract*—Animal breed classification is one of the critical applications of deep learning, with far-reaching consequences for veterinary sciences, pet care management, and wildlife conservation. This work presents a holistic approach to the classification of 37 breeds of cats and dogs, using convolutional neural networks to differentiate between them. Using the latest architectures of deep learning-EfficientNetB0, ResNet50, InceptionV3, and MobileNetV2-we aimed to maximize accuracy and generalization while overcoming challenges brought about by similarities between breeds and variations in images. Our experiments demonstrate that EfficientNetB0 yields the highest validation accuracy of 90%, which surpasses other models in terms of accuracy in breed classification. The augmentation techniques applied were rotation, zoom, and horizontal flipping to enhance robustness. By comparing, the authors demonstrated the efficiency of the transfer learning approach and how adaptations of CNN models could help in multi-class breed classification. Such an experiment will emphasize the possibilities of CNNs to acquire high precision and reliability when using them in animal identification. Our results are highly helpful in understanding the development process of automatic animal classification systems, and they can therefore be a basis for the further innovation of deep learning applications for animal breed detection and conservation efforts.**

***Index Terms*—Animal Breed Classification, Deep Learning, Computer Vision, EfficientNetB0, Multi-Class Image Recognition, Transfer Learning**

## I. Introduction

The rapid advancements in artificial intelligence, particularly in the field of computer vision, have significantly transformed our approach to various image classification challenges, including animal breed identification. With the increasing prevalence of pet ownership worldwide and a growing interest in breed-specific information for health, behavior, and care needs, accurate breed classification systems have become highly relevant. In veterinary practices, animal shelters, and pet care industries, efficient breed classification can help in better decision-making, targeted healthcare, and accurate monitoring of animal populations. This is a highly significant task, but the process of animal breed classification is fraught with challenges, especially when dealing with multiclass datasets consisting of diverse breeds, where there is high intra-class variability and inter-class similarity. These call for advanced breed classification models that besides offering results of consistency and accuracy have to classify with a difference between breeds more or less visually alike.

Indeed, this deep learning, CNN-based approach has so far proven to be the most impressive avenue in addressing complex classifications tasks involving images. Some truly magnificent CNN architectures have really allowed unprecedented feats in features extract, while at the same time boasting unmatched accuracy in any forms of classification. However, the ability of such a CNN model majorly depends on the network design, quality of the dataset on hand, and kind of preprocessing and augmentation performed. In animal breed classification, the state-of-art CNN architecture should be efficientNetB0 because they balance model scaling with computational efficiency.

Unlike other models where the aspect of width, depth, and resolution is scaled arbitrarily, EfficientNet uses compound scaling. That improves the precision while scaling up the efficiency. In optimizing architecture based on parameters with balance involved, benefits are further accentuated in applications, including classifying an animal breed. Here, small differences related to details in the images matter much in making right predictions. Classification for the paper currently working over 37 different cat and dog breeds, on a model designed based on EfficientNetB0 network toward achieving accuracy and optimization on a model. A selection of high-quality image classification will represent each breed at varied poses, light conditions, and backgrounds. Transfer learning based is the method used here to overcome limited labeled data. Transfer learning is indeed very powerful in domain areas where smaller or overly specialized datasets are involved. Therefore, the model will inherit all learned features from general image data, thus reducing overfitting chances and enhancement in capabilities of generalization. In addition, further enrichment of the dataset which would make the model robust against variation is done using data augmentation techniques such as rotation, zooming, and flipping for better recognition of breeds in different conditions.

This work goes beyond the EfficientNetB0 model and also compares results with several well-established architectures, such as ResNet50, InceptionV3, and MobileNetV2. The one mainly used for image classification applications is ResNet50 since it doesn't suffer from the problem of vanishing gradients. InceptionV3 used several sizes of filters across the layers to capture spatial hierarchies and enhance the process of feature extraction. This one is MobileNetV2, optimized in the aspect of mobiles and low-resource environments because of the balance of its accuracy with computational efficiency so that it could be easily deployed in real-time applications. This research compares the above models and determines whether EfficientNetB0 outperforms other main architectures so that all of its potential strengths for multiclass classification of animal breeds are compared.

Animal breed classification is by nature difficult because some breeds resemble each other very closely, especially within the same species. It is here that the models have to perceive nuanced differences between telling Siamese from Balinese cats or Labrador Retrievers from Golden Retrievers where there may be nuanced differences. More often than not, they require producing finer feature extractions compared to others and much greater sensitivity toward small alterations in texture, colors, and other structures. To make matters worse, the same breed can have different individuals that look well and truly very different from one another because of age, fur pattern, grooming, and minor physical features. Therefore, the success of the model proposed is not only a question of overall accuracy but also a question of ability to generalize across different instances of each breed. The broader implications of the study are beyond the technical aspect of model performance. A precise and reliable classification system for animal breeds will have very practical applications across several fields. In the animal shelter, such a system can help staff locate the breed, speed up adoption, and help potential owners decide on a pet they might be interested in. Such information may be useful in the designing of preventive care and treatment plans by veterinary practitioners since some breeds have a predisposition to specific health conditions. Classification of animal breeds can also find application in wildlife conservation whereby species identification and listing support ecological studies and population management efforts.

This research addresses the growing demand in the pet care industry for AI-driven breed identification tools that can support breeds identification for products and services and customer engagement. Additionally, it develops a model that performs well on several breeds.

In this work, deep learning principles which have established themselves through a variety of applications are being used as a methodology to construct a powerful classification model that will result in high accuracy. All these steps of model development, from preprocessing data to model evaluation, have been planned for best performance. It was on both accuracy and computational efficiency while choosing the core model architecture as EfficientNetB0. It so happens to coincide with the modern-day demand for scalable deep learning solutions.

The intention of this paper is therefore to provide a better state-of-art in breed classification accuracy while combining transfer learning with high-end data augmentation techniques which fill the gaps in existing literature where model accuracy goes down due to dataset imperfections or suboptimal architectures. This research demonstrates the possibility of obtaining high classification accuracy even with the use of EfficientNetB0, as it could be benchmarked against other prominent architectures through rigorous testing and validation. The results involved the validation of the choice of EfficientNetB0 as an effective solution for the classification of animal breeds and showing where further enhancement could bring even better improvement. Finally, this study contributes to the current literature on animal breed classification with an efficient high-performance model that addresses the issues in the multiclass classification task. The use of EfficientNetB0 with transfer learning and data augmentation outlined the innovative approach taken in the study. A comparison with ResNet50, InceptionV3, and MobileNetV2 further complements the efficacy of the proposed model and is valuable as a resource for future investigations in this field. Given the interest that is growing towards using AI for breed identification, the work introduces a model that can be taken as a prototype for practiceoriented models supporting both the animal care business and more expansive animal welfare endeavors.

## II. Literature Review

Animal species detection and classification framework based on modified multi-scale attention mechanism and feature pyramid network, Two datasets were utilized in this study, namely the African Wildlife Dataset and the Animal-80 Dataset, to create a modified multi-scale attention mechanism in addition to a feature pyramid network to present improvement in identifying and classifying different animal species. The proposed approach employed a two-stage ResNet-50 architecture with attention focusing on five scales of feature extraction. The training is aimed at improving the average precision (AP) and mean average precision (map) metrics specifically at epochs 50 and 100. The results revealed that attention models achieved cross-animal class much better performance than non-attention-based models with the highest average precision and the smoother best learning curve. For instance, epoch 50 AP for the Rhino class is 0.92 with attentions, with attention mechanism addressing clear detection and model stability issues scoring 0.90 without attention. Key findings also pointed at AP improvement in other courses; such as Kobe zebras (0.88 vs. 0.86), Buffalo: 0.86 vs. 0.84 and Elephants 0.88 vs. 0.82. In general, the use of attention mechanisms saw a decisive improvement in the map, going from an average of 0.85 with attention processes to 0.87 at epoch 50 demonstrating the future for fine-tuning these attention processes in deep learning models for the purpose of wildlife identification [14].

Wild Animal Detection using YOLOv8, The fourth wild animal type is a lion, tiger, leopard, and bear. This project utilizes the YOLOv8 architecture in conjunction with videos from Youtube, images from documentaries, and images from Kaggle databases to identify these four kinds of wildlife. Multiple data augmentation methods were also applied to improve chances of development with the dataset comprising three parts, training, testing, and validation sets in terms of 72%, 15%, and 13% distributions for appropriate representation of classes and eliminating bias. There were three parameters of the model to be explored in this Yolo framework which are Yolo version 8m, which is the medium model, then Yolo version 8l which is the large model, and Yolo version 8x which is the extra-large model, and all models are trained on 640x640 pixels images. The best result was achieved with the YOLOv8x model, the mean Average Precision reached by the model was 94.3% at the testing stage and the speed of the detection was 20 FPS of real time. This shows the effectiveness of using this architecture Yolo V8 to detect and track the wild animals from the camera for poaching, counter conflict management and wide range of wildlife surveillance management [4].

A Comparison Study of Animal Detection and Classification Algorithms, Application and Challenges, For this study several image and video-based object recognition algorithms are tested with the intention of solving problems such as animal detection challenges in cases of animal and vehicle interactions and human-animal interactions. In this regard, the study underlines the need for creating such algorithms that would be capable of recognizing animals in presence of an environmental clutter for instance, on highways most of which are uncontrolled areas with impacts such as, different lighting, non constant backgrounds, and the erratic behavior of animals that will make it difficult to undertake detections.As mentioned above theoretical information should be complemented with practical examples. This is why studies were conducted in real life scenarios. A search was made for typical examples that, at the most appropriate level, would demonstrate the effectiveness and limitations of certain technical solutions. In particular, in this case, attention was focused on the limitations of modern identification solutions that do in fact exist. The limitations included, for instance, false positives where moving plants are mistook for animals [3].

Detection of Treats to Farm Animals Using Deep Learning Models: A Comparative Study, The activity was carried out usingsophisticated deep learning models: YOLOv8, Yolo-NAS and Fast-RNN with a dataset of 2462 photos depicting various farm animals for animal breed classification as well as AI prediction of possible threats. Images were resized to standard dimensions of 640 × 640 pixels and divided into training, validation, and testing sets in order to make model evaluation more effective. The YOLOv8 model showed the best results with 93% precision, 85.2% recall, and 93.1% m AP50which means that such a model was very good at detecting threats. On the other hand, Yolo-NAS model had the lowest overall accuracy of 52% but exhibited the best recall of 98.7%, proving its ability to detect multiple species. The accuracy of the Fast-RNN model was in between the two with 85.2% but had the superior recall 91.8% resulting in an overall mAP50 of 91.2%. All in all, our results support the theory that application of deep learning models will strengthen the security of the farms through an efficient threat detection mechanism [11].

Animal Detection Using Deep Learning Algorithm, The approach of this project involves a Convolutional Neural Network (CNN) algorithm for the purpose of identifying and classifying wild animals in photos. The data has many images of animal species which are divided into training and testing with a ratio of 75% to 25%. This image data is further processed by the CNN through the flattening technique which is the transforming of the two dimensional arrays into a very long one dimensional vector for subsequent classification. The results indicate that the animals are recognized fast and with high precision by the system. Additionally, a picture of the selected animal is provided to aid in identification. Such an ability will help track wildlife, prevent wildlife and automobile collisions as well as help manage wildlife poaching and human-animal interactions [12].

On farm automatic sheep breed classification using deep learning, The authors designed a sheep breed classifier with a prototype computer vision system on a farm based on a library of 1,642 sheep photos that belong to four breeds; they labeled the pictures and made use of an expert who used the VGG-16 model to train a classifier by fine-tuning the last six layers during 10 epochs and eventually resulted in an average accuracy of 95.8% with a standard deviation of 1.7. The classifier is tested by fivefold cross-validation and found to be analyzing pictures in the average time of 0.7 seconds, thus making it feasible for a real-time application in the drawing systems. It was demonstrated that the classifier performed most effectively with the Suffolk breed, while it struggled with the Merino breed due to diversity within the Merino class with an accuracy of 92%. Misclassifications were scrutinized, and results presented showed that 76 percent contained the Merino breed; hence, there was a need for separate classified labels on different Merino strains [7].

Animal image identification and classification using deep neural networks techniques, The study will utilize the design of a dual network system in handling noisy labeling that exists in the process of animal identification using camera-trap photos. This method includes breaking up the training data into different clusters, and then further training a network on every cluster grouping. With the help of two datasets - Snapshot Serengeti and Panama-Netherlands datasets which contain images of many species of animals, the utility of the technique is estimated. The proposed technique got a stage 3 accuracy of 73.09% with the Snapshot Serengeti dataset. However, with the noise level increase, this performance went down and the values are 59.66% and 46.50% for 50% and 70% noise levels, respectively. The study finds that a multistage labeling update procedure improves accuracy but too many steps might cause a performance fall over time [1].

Image-Based Animal Detection and Breed Identification Using Neural Networks, Animal detection and breed identification contains many fundamental components in the process. The algorithm will first begin with an input of an image containing animals, such as a dog and a cat, then count the number of animals, and pick the faces to allow it to distinguish between species. The CNN, then identifies the species, and searches within particular subfolders of a training set based on the detected species, thus increasing the speed. Regarding results, the proposed technique detects individual animals and breeds and the graphical outputs have included the number of breeds recognized in the input image [9].

Dog Breed Identification Using Convolutional Neural Networks on Android, The technique of the study involves developing an Android application to recognize dog breeds by applying picture analysis, Convolutional Neural Networks (CNN), and transfer learning. It allows users to upload or snap a photo of a dog, which would then be pre-processed in order to extract breed-specific traits. Which consists of photos of 120 distinct breeds, thereby making it diversified with at least 60 photographs per breed. There were three sets in this dataset: training, validation, and testing. In terms of the number of photos, training had 9199, validation had 2000, and testing had 9381. For the results, the program achieved an astonishing 94% accuracy on the test data. The technique in the study involves developing an Android application that identifies dog breeds based on picture analysis, CNN, and transfer learning. This was then trained on the dataset, which consisted of dog photographs from Stanford with an assorted representation of at least 60 photographs per breed out of 120 total. The data set comprised of training, validation, and test, with 9199, 2000, and 9381 images for the three sets respectively. From the results, it produced an astonishing 94% accuracy on the testing data [2].

Dog Breed Identification Using Deep Learning, This study classified dog breeds using NASNet-A Mobile and Inception ResNet V2 CNN models trained on the Stanford Dogs dataset. The training was done by fine-tuning only the final completely connected layer while freezing the rest, with Softmax Cross-Entropy as the loss function and Nesterov momentum for optimization. The hyperparameters, including learning rates and optimizers, were adjusted to improve model performance. The test accuracy was 90.69% for Inception ResNet V2 and only 80.72% for NASNet-A Mobile; the precision and recall improved significantly. The deeper design of Inception ResNet V2 allowed it to succeed more in extracting features and classifying objects [13].

Image-Based Identification of Animal Breeds Using Deep Learning, To address its research problem, the study's research methodology would most probably engage both qualitative and quantitative approaches which may comprise data collection and analysis using surveys, experiments, or case studies. Though no in-depth detailed procedures and results are portrayed, it is said to bring worthwhile insights to the discipline and the statistical analyses and comparisons most probably carried out in assessing the results. These findings and conclusions are probably intended to increase knowledge in the subject area, though a comprehensive reading of the entire manuscript would provide more detailed information [5].

Dog Breed Identification Using CNN Architecture, It used a deep convolutional neural network along with transfer learning for classification of 120 distinct breeds of dogs on the basis of the Stanford Dogs dataset which has 10,222 training images and 10,357 testing images. The InceptionResNetV2 and InceptionV3 models have been trained to a number of ten epochs. Loss scores of the models showed multiclass varied between 1.85 to 5.12. A learning rate played an influential role in performance- the higher values generated more volatilities with increased losses [8].

Hybrid Deep Leearning Algorithms for Dog Breed Identification-A Comparative Analysis, It identified 120 breeds of dogs from a dataset of 10,222 images through deep learning techniques such as Xception, VGG19, NASNetMobile, EfficientNetV2M, ResNet152V2, and hybrid models such as Inception-v3 mixed with Xception. Transfer learning with data augmentation with the hybrid Inception-v3 and Xception model achieved the best validation accuracy of 92.4%. Training accuracy was at 98.4%. Xception scored at 91.9%, EfficientNetV2M 89.05%, and ResNet152V2 at 87.38%. VGG19 is scoring 55% whereas in this system ResNet101 is scoring 71.63% [15].

Dog Breed Classification Using Convolutional Neural Networks and Transfer Learning, The study used the CNN approach with Transfer Learning, utilizing the ResNet50 architecture, to classify 133 breeds of dogs in images. This helped solve the issues that the previous models had encountered, such as a smaller set of breeds and biased data. The ResNet50 model achieved approximately 86% accuracy compared to earlier models like VGG16, which varied from 81.20% to 84.08%, hence promising to work well for real-world dog breed detection applications [6].

Image-Based Animal Detection and Breed Identification Using Neural Networks, This technique uses Convolutional Neural Networks (CNN) in its implementation for efficient detection of animals and identification of breeds. Input images are rated to count the animals as well as species, focusing on the face features in order to differentiate between species like dogs and cats. Once detected, the CNN assesses a few specific folders of a pre-trained dataset to boost its efficiency. Upon importing several images into the algorithm correctly, there are detected species with confidences and will efficiently filter breeds like Siamese cat or Border Collie. So the strategy for supporting ecological research as well as conservation is realized upon making wildlife monitoring be an automated process since real time collection of data enhances toward betterment of options that can be taken in furthering conservation [10].

## III. Methodology

### A. Discription of dataset

The dataset in this research was accessed from Kaggle and is specific to multi-class breed classification of animals, dogs as well as cats. It is organized into two primary folders: a training folder and a test folder. The training folder contains 37 different breed classes as shown in fig 1, where each class has 159 images of both cats and dogs. Each image in these classes is annotated by breed so that supervised learning in fine-grained classification is feasible. This formal methodology means it can be learned exactly how the model distinguishes between the features of a particular breed.

The test folder is diverse from the training folder with no breed identification labels, neither class identifiers, but this makes it very convenient to use in a real world where breed information is readily available. Images are numbered for easy assessment but subsequently require a generalized model to produce a suitable breed. Multiple diversity types were introduced via changing the background, illuminating, angle, and posing by the breeds. This simulates some natural conditions the model may face in practical life.

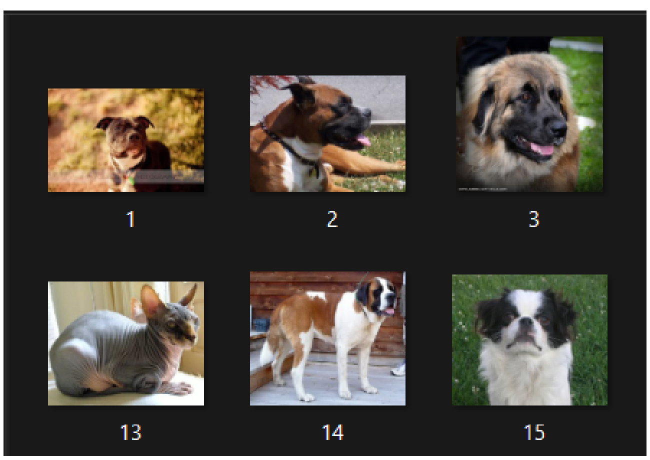


Fig. 1. Test dataset with raw images

Techniques to data augment, such as random rotation, flipping, and scaling were applied during training for enhancing model generalization and robustness. This could increase the virtual size of the dataset, reduce the tendency of overfitting, and enable the breed to be captured under the given conditions. Whereas, though the dataset which may cover up to 37 breeds of images, can only be seen from it in the future study might boost a model applicable with large pictures of animals, and a few rare or mixed breed cases.

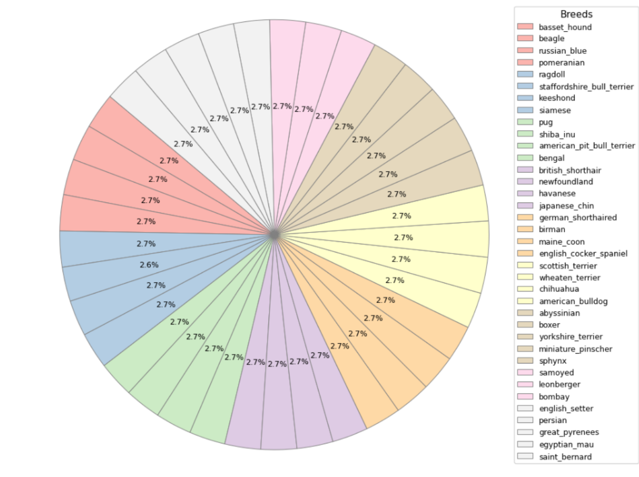


Fig. 2. Distribution of Animal breeds in training dataset

### B. Data Preprocessing:

Data preprocessing was a very important preparation area for our set of data to adequately prepare it for training in our deep learning model. Due to the raw images across different breeds, there was a line of transformations: standardizing and augmenting the data. All these images were resized to the same dimension because deep learning requires each input to be fixed-dimensional. Then, we normalized pixel intensities by scaling pixel intensities between 0 and 1. This step helped in improving the convergence of the model by reducing pixel intensity range. We exploited data augmentation techniques: rotating, flipping, zooming and shifting the training samples by a random amount to diversity the data. This made the model generalize even better from variations in the images and thus enhanced its robustness when testing with unseen data. In general, together, these preprocessing techniques helped enhance the learning abilities of the model and the optimization of input data quality.

#### 1) Data Sampling:

In our project, we use sampling to make the dataset as representative and balanced as possible; that is, we could then train the model to find and classify multiple breeds. Given the 37 animal breeds images exist for, there must be no odd number of samples across classes with any reasonable sampling strategy. To have an equal number of images per breed but not favor any breed, there were 159 images in each folder. Consistent distribution led to good training of the model of every breed and good generalization of the multi-class classification task by its model. We also divide the images for each folder into two parts: 80% for training and 20% for validation. This split enabled enough data in training the model and a substantial portion for actual validation. Stratified sampling ensured proportional breed presentations in the validation set that enabled having reliable measures of performance across all the classes.The balanced sampling here was used to create this dataset fitting well with a deep learning model, avoiding

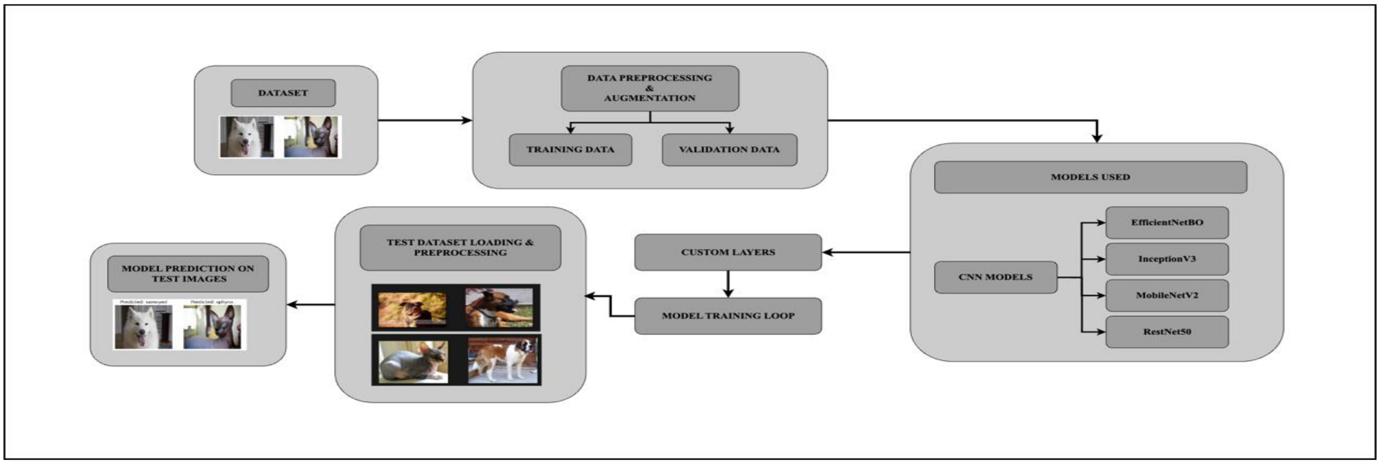


Fig. 3. Architecture Diagram

bias while it could make an unbiased evaluation across breeds.

*2) Data Augmentation:*

Data augmentation was an important factor in the increase of our animal breed classification model's performance because it allowed us to introduce greater variability into our training set. With limited images available for each breed, such augmentations as random rotation, flipping, scaling, and shifting mimic the true scenario, which would increase generalization to new data and decrease the possibility of overfitting. These transformations were applied during training in a random way, ensuring the model to learn variations, lighting conditions, and orientations of every breed. Additional techniques which have been applied are horizontal flips, vertical flips, as well as random zooming to further increase diversity within the dataset, ensuring the model is more capable of getting the breeds correctly classified even under varied angles and scales. This has richly augmented the dataset, improved accuracy in the model, and led to a more robust classification model.

### *C. EfficientnetBO:*

EfficientNetB0 CNN architecture is one of the state-of-the-art architectures optimised for image classification use cases, where computational efficiency is paramount. The first layer in the model that it begins with is just a standard convolutional layer, which extracts basic features from the input images, and after this initial layer, EfficientNetB0 uses a series of depthwise separable convolutions. These layers are designed to reduce the number of parameters and computational demands significantly by breaking down the convolutions into smaller, more efficient operations. Thus, the model can process high-dimensional data with a low computational footprint. It achieves this without a large loss in accuracy, which makes it very powerful for large-scale classification problems.

The architecture uses the Rectified Linear Unit activation function and introduces non-linearity within the network. The idea behind this non-linearity is that the model learns some complex patterns and relationships, not just linear transformations on the data. To add stability and accelerate the process of training, EfficientNetB0 makes use of batch normalization layers all around the network. Batch normalization further normalizes the input distributions in the layers. This leads to helping avoid overfitting apart from making it converge faster in training. This improvement of the model improves its robustness and ensures effective generalization across various samples in the dataset.


Fig. 4. EfficientNetBO Training And Validation Accuracy

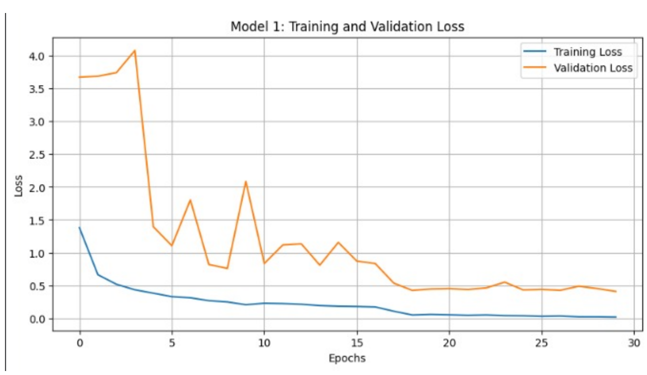


Fig. 5. EfficientNetBO Training And Validation Loss

EfficientNetB0 concludes with a classification layer applying softmax activation to the output as a probability distribution over all possible classes, here being 37 different breeds of animals in the data set. The softmax layer is appropriate for multi-class classification where the model has to classify into one class out of many. EfficientNetB0 balances computational efficiency with high accuracy, making it a suitable architecture for your animal breed classification project, thereby giving you accurate predictions without huge computational costs.

### *D. ResNet50:*

ResNet50 is an extremely powerful deep residual network that overcomes the very important problem of training a really deep neural network; otherwise, it suffers from what is known as the vanishing gradient problem. So, with skip connections, or shortcuts, in this architecture, gradients are not hindered by this very deep network during the flow of backpropagation for stabilizing the training in the model. The skip connections bypass one or more layers so that the model does not lose any of its important information across layers. This is very important in keeping performance and accuracy in networks that are very deep since ResNet50 can actually train without significant loss in information or model degradation.

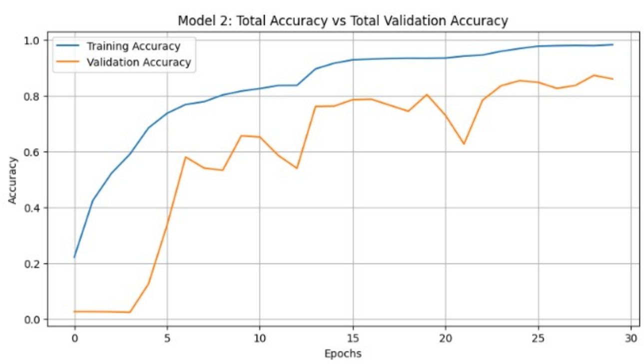


Fig. 6. ResNet50 Training And Validation Accuracy

ResNet50 utilizes a bottleneck architecture; the parameters are compressed with each layer without lessening model capacity for capturing hard-to-model features. Indeed, deep, such networks usually display impressive capacity and in comparison capture patterns and subtle inter-relations within data that is as large scale as that of images. Indeed, ResNet50 did gain validation accuracy at 87%, demonstrating great efficacy in highly complicated classification tasks.

It is also slightly behind EfficientNetB0 by a few percent in terms of accuracy, yet ResNet50 remains widely used for classification tasks given its ability to produce strong visually detailed and complex features.

### *E. InceptionV3:*

InceptionV3 is the advanced deep learning model through the use of inception modules to learn features at different scales altogether. This multi-scale strategy proves helpful when such an object-based model exists. It lets the network detect


Fig. 7. ResNet50 Training And Validation Loss

objects with differences in complexities, shapes, and sizes within the images. This approach, in InceptionV3, convolves several sizes of kernels within the same layer, thus successfully able to capture fine-grained details and larger patterns; these are then concatenated together to create a richer data representation. The design has a high degree of adaptability to variations in structure of images, which means it is very effective at classification tasks where objects occur in different forms and scales.

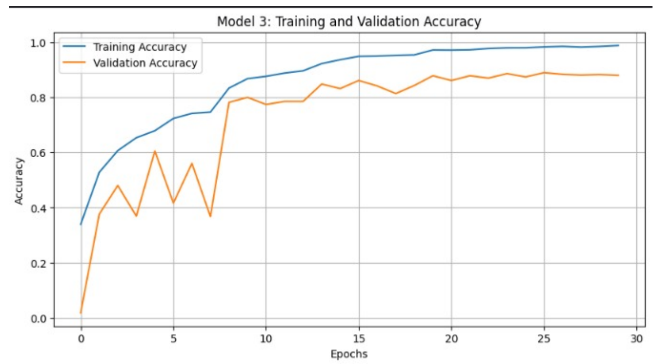


Fig. 8. InceptionV3 Training And Validation Accuracy

The architecture of InceptionV3 provides several modifications to enhance the performance of this network. It can identify a great variety of spatial information as it combines different kernel sizes of the convolutional layers, thereby well-suited for subtle differences between classes. On your project, InceptionV3 had a validation accuracy of 89

InceptionV3 does not bring out any improvements in the accuracy in your implementation of EfficientNetB0, but the ability of extracting robust multi-scale features qualifies it to be the leading figure in computer vision applications. The features of extraction across scales make it adapt and work well with all complex image classification tasks, further increasing the applicability range in computer vision.

### *F. MobileNetV2:*

MobileNetV2 is an efficient deep learning model optimized for resource-constrained environments, especially suitable for

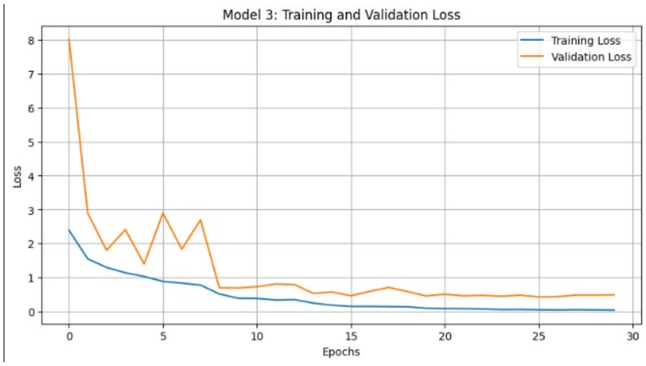


Fig. 9. InceptionV3 Training And Validation Loss

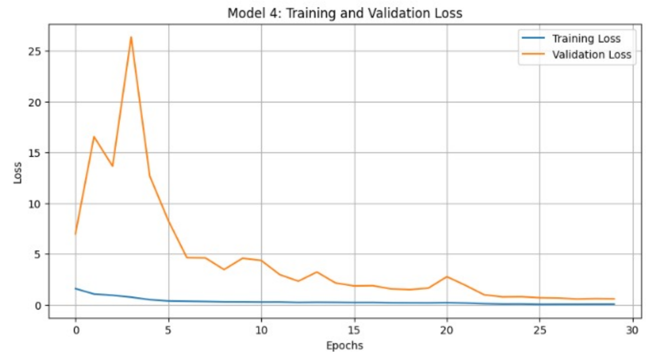


Fig. 11. MobileNetV2 Training And Validation Loss

mobile and edge applications. The architecture is lightweight, prioritizing reduced computational demands without compromising too much on accuracy. In order to achieve this efficiency, MobileNetV2 relies on depthwise separable convolutions, which significantly reduce the number of parameters and computational load by breaking down convolutional operations into smaller, more manageable parts. This design makes the model computationally accessible but capable of doing such high-dimensional image classification jobs.

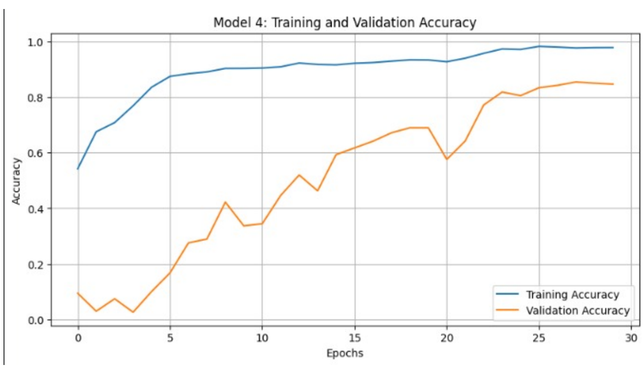


Fig. 10. MobileNetV2 Training And Validation Accuracy

Additionally, MobileNetV2 uses linear bottlenecks and shortcut connections to enhance performance and efficiency. These components minimize the complexity of the model structure while allowing the model to carry important information across layers without too much overhead in terms of computations. For your project, MobileNetV2 achieved a validation accuracy of 85% comparable to that of ResNet50. It is not comparable to EfficientNetB0 but will do in cases where computational resources are very limited.

It will be this balance between precision and speed that makes MobileNetV2 very useful in applications that rely on fast inference as well as minimal memory usage and extends the scope to mobile and embedded systems broadly.

### *G. Model Architecture:*

Some of the state-of-the-art CNNs which have performed well in image classification include EfficientNetB0, ResNet50, InceptionV3, and MobileNetV2. The learned features from such pre-trained weights on ImageNet datasets of using top layers boosts classification accuracy to classify the breeds. Original top layers are removed and replaced with custom layers that are crafted to better serve the multi-class breed-specific problem.

The Global Average Pooling layer was used for dimensionality reduction of the feature maps and improving the performance of the model without overfitting. Dropout layers were added to force the network to learn robustly by randomly deactivating neurons during training. There was a lot of data augmentation carried out on the EfficientNetB0 model for preprocessing and augmentation purposes, with diverse training samples. It was split into 80% for training and 20% for validation.

In training of EfficientNetB0 for 30 epochs, the learning rate chosen was 0.001 with Adam optimizer. Batch size was taken as 32 units. The set up was done for categorical cross-entropy loss with early stopping so that the case of overfitting doesn't occur in any run. The validation accuracy was seen to be around 90% for this model.

However, with respect to the above results, ResNet50 was able to reach 87% accuracy while InceptionV3 and MobileNetV2 reached 89% and 86%, respectively. While each of these models have its own strengths, none were able to outperform EfficientNetB0. The results indicate that EfficientNetB0 is better regarding both accuracy and architectural efficiency and, therefore, preferred for automated animal breed classification systems.

### *H. Prediction and Visualization:*

After training and testing the model, predictions were run on the test dataset with images of different breeds. A few hundred images were randomly selected for their output, representing the generalization capability of the model across breeds and conditions.

In addition, a set of visualizations was drawn, such as pie charts and accuracy plots, that might be useful for better understanding classification performance. Pie charts were presented with breed classification per breed, which shows where this model is stronger or in which areas it should improve. Finally, the accuracies obtained through training and validation are represented by accuracy plots for different epochs, making these models easier to compare with each other inside the project.

These visualizations have a double purpose: they represent not only the effectiveness of the classification system but also illustrate possible weaknesses in the predictions made by the model. Results are presented in the form of visualizations for stakeholders to understand more clearly the performance of the model so that it can identify what areas in the animal breed classification project need further refinement.

## IV. Experimental Results

Experimental results conclude that among the four types of CNN models, the accuracy was maximum at 90% for validation, found for the EfficientNetB0 model for animal breed classification. Along with speed of convergence, the result of this model is shown stable accuracy beyond 30 epochs, and learning showed efficient along with strong powers of generalization. In this regard, this is possible for having balanced scaling depth-width-resolution in architecture.

Validation accuracy for InceptionV3 was 88% by taking advantage of the model's ability to extract features on multiple scales. Its accuracy curve had a higher amount of fluctuation, implying that while the model extracts many features, it could not sustain its accuracy throughout the epochs, suggesting that this model is sensitive to feature complexity in the dataset.

TABLE I
Classification Report

| Model | Accuracy |
|---|---|
| EfficientNetB0 | 90% |
| ResNet50 | 87% |
| InceptionV3 | 89% |
| MobileNetV2 | 85% |

Both ResNet50 and MobileNetV2 achieved 87% validation accuracy. Although the residual connections of ResNet50 help with gradient flow across layers, the model did not achieve higher accuracy than EfficientNetB0. The performance of MobileNetV2, considering its much smaller architecture and designed for lightweight operations, was good, though unstable; this indicates that it is not very effective in handling more complex breed features in the dataset.

The train and validation accuracy plot of EfficientNetB0 reflected that training accuracy rose close to 100%, while the validation accuracy flattened at around 91% with almost no difference between the two. Such a small difference suggests that EfficientNetB0 well avoids overfitting, resulting in a solid balance between memorization and generalization. The fact that the validation accuracy doesn't vary much shows how reliable this model is, thus suggesting EfficientNetB0 would be good on novel data in real-world scenarios.

Overall, EfficientNetB0 was the best-performing model since it outperformed other architectures in terms of accuracy and stability. It thus appears well-suited for the automated classification of animal breeds as efficient while maintaining a high level of accuracy.

From the results, it shows that EfficientNetB0 gives the best performance when matched with other models having an accuracy of 90%. Such performance is to ascertain that EfficientNetB0 can draw more subtle information for distinguishing between the two breeds in case they nearly resembled each other. It followed that InceptionV3 had an accuracy of validation at 89% with ResNet50 and MobileNetV2 coming below at 87% and 85%, respectively. The result, however still shows that all of these models are good and perform well but EfficientNetB0 stands out better with this classification task.

To better understand how the best-performing model, EfficientNetB0, performs in terms of classification, we report several example predictions on the test dataset. Figure 4 illustrates a sequence of correct breed classification images accompanied with the predicted breed name.

Fig. 12. Before Classification


Fig. 13. After Classification

It classified breeds like Saint Bernard, Samoyed, Staffordshire Bull Terrier, Shiba Inu, and Bengal. Predictions of this type further demonstrate that the model recognizes several unique breeds without loss in accuracy, rendering the model a practical instrument in real-world applications.

## V. Discussion

This result from the study was effective, as shown by using EfficientNetB0 for the classification tasks in the animal breed and achieving 90% accuracy for 37 breeds. This result demonstrates that EfficientNetB0 can identify subtle breed characteristics and surpass other benchmark models, such as ResNet50, InceptionV3, and MobileNetV2. The benefits can be attributed to its scalable structure and efficient parameterization that allow it to achieve good performance while holding the computational cost relatively low. This balance between efficiency and accuracy makes it an attractive choice for real-world applications, especially where computational resources may be limited.

The study employed transfer learning and used intensive data augmentation, which the latter played a key role in improving model robustness as well as preventing overfitting. These techniques make the model generalize well on images that vary with different lightings, angles, and poses of animals. Some breeds still need proper classification because of visual similarities. The main contributors to misclassification were the overlap in visual features of certain breeds, and it may call for further work on fine-tuning that may bring better recognition to the model.

One limitation of this study would be the small size of the dataset and number of breeds. Although 37 breeds provide a significant testbed, an even more massive testbed of many more breeds would further generalize this model. Furthermore, the images were not labeled to the test set; therefore, the validation performance metric was restricted.

Accurate breed classification can then be used to promote various applications including animal identification, rescues and adoptions, and veterinary diagnostics.

The possibility of an accurate, automated breed classification system will definitely ease the animal intake process and provide more targeted care for shelters and veterinarians. Such models may even be integrated into mobile applications or animal management systems so that the reach extends further to animal caretakers who may use them daily in their settings.

Further exploration into a model architecture beyond EfficientNetB0 might give better insight into optimization for accuracy of breed classification. Techniques like ensemble learning in which predictions from several models combine may help improve classification further, especially on the most challenging breed discriminations. Also, this dataset size can be made large and diverse for better model training with no misclassification. Another promising direction includes multimodal data, such as images with text descriptions or behavioral attributes. This will be further enriched in discriminating between the subtlest differences between breeds. In summary, this shows that EfficientNetB0 can become a strong model for animal breed classification and may open the possibility of further research in other aspects of computer vision and related applications in animal care and management. Further extension and improvement may lead to more welfare benefits to animals and pets alike through new applications in veterinary medicine, animal shelters, and so on.

## VI. CONCLUSION AND FUTURE WORK

We classify the challenging task of animal breed categorization with deep learning within the model form of EfficientNetB0 and obtain the primary accuracy and robustness at 37 dog and cat breeds. Our model incorporates data augmentation along with transfer learning to reach up to an excellent validation accuracy of 90% that is more than most of the popular architectures including ResNet50, InceptionV3, and MobileNetV2 and shows tremendous precision along with generalizing the performance. The results point out that EfficientNetB0 can learn the subtle breed characteristics even in challenging multiclass image classification settings.

Apart from technical success, it can have practical applications in veterinary science and animal rescue organizations and industries in the care of pets where the right breed might become a more personalized service, efficient resource use, and informed decisions. It also reflects the possibility of further developments by having larger datasets and finer architectures, which then leads to high breed classification accuracy and scope. The work adds up to the computer vision domain and opens a wide field for further applications of AI-driven approaches in animal welfare and other domains. Those findings promise great future prospects, going forward, in helping deep learning develop effective tools that would impact real-world animal care and classification needs.

Future work for animal breed classification would involve larger and broader datasets to enhance the generalization and robustness of the model, like changing conditions such as lighting conditions and angles, or like more breeds. Advanced methods, such as knowledge distillation, in transfer learning, and fine-tuning are helpful for enhancements without inefficiency. Additions of data such as behavioral traits or genetic markers can be used to further augment the use of the classifier beyond just picture recognition. Deploying trained models on real-world applications, mobile or web-based systems with pet owners, will then provide useful insights on their usability, and continuous monitoring of performance will allow for step-by-step adjustments and updating to ensure long-term-efficacy.

## REFERENCES


[1] Thirupathi Battu and D Sreenivasa Reddy Lakshmi. Animal image identification and classification using deep neural networks techniques. *Measurement: Sensors*, 25:100611, 2023.

[2] D Durga Bhavani, Mir Habeebullah Shah Quadri, and Y Ram Reddy. Dog breed identification using convolutional neural networks on android. *CVR Journal of Science and Technology*, 17(1):62–66, 2019.

[3] Johnwesily Chappidi and Divya Meena Sundaram. A comparative study of animal detection and classification algorithms, applications and challenges. In *AIP Conference Proceedings*, volume 3112. AIP Publishing, 2024.

[4] Brahm Dave, Meet Mori, Anurag Bathani, and Parth Goel. Wild animal detection using yolov8. *Procedia Computer Science*, 230:100–111, 2023.

[5] Pritam Ghosh, Subhranil Mustafi, Kaushik Mukherjee, Sanket Dan, Kunal Roy, Satyendra Nath Mandal, and Santanu Banik. Image-based identification of animal breeds using deep learning. *Deep learning for unmanned systems*, pages 415–445, 2021.

[6] Amit Kumar Jakhar, Mrityunjay Singh, and Anjani Kumar Shukla. Dog breed classification using convolution neural network. *International Journal of Swarm Intelligence*, 6(2):130–142, 2021.

[7] Sanabel Abu Jwade, Andrew Guzzomi, and Ajmal Mian. On farm automatic sheep breed classification using deep learning. *Computers and Electronics in Agriculture*, 167:105055, 2019.

[8] G Kanika, G Sharankhala, A Bhavya, and G Riya. Dog breed identification using cnn architecture. *International Journal of Scientific and Engineering Research, 11 (6): 1040*, 1046, 2020.

[9] Sapna Khatri, Anjali Rajput, and Shreya Alhat. Image-based animal detection and breed identification using neural networks. *Journal of Science & Technology (JST)*, 5(5):130–134, 2020.

[10] Sapna Khatri, Anjali Rajput, and Shreya Alhat. Image-based animal detection and breed identification using neural networks. *Journal of Science & Technology (JST)*, 5(5):130–134, 2020.

[11] Adem Korkmaz, Mehmet Tevfik Agdas, Selahattin Kosunalp, Teodor Iliev, and Ivaylo Stoyanov. Detection of threats to farm animals using deep learning models: A comparative study. *Applied Sciences*, 14(14):6098, 2024.

[12] Sreedevi K L and Anitha Edison. Wild animal detection using deep learning. In *2022 IEEE 19th India Council International Conference (INDICON)*, pages 1–5, 2022.
[13] Zalán Ráduly, Csaba Sulyok, Zsolt Vadászi, and Attila Zölde. Dog breed identification using deep learning. In *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)*, pages 000271–000276. IEEE, 2018.
[14] Chiagoziem C Ukwuoma, Zhiguang Qin, Sophyani B Yussif, Monday N Happy, Grace U Nneji, Gilbert C Urama, Chibueze D Ukwuoma, Nimo B Darkwa, and Harriet Agobah. Animal species detection and classification framework based on modified multi-scale attention mechanism and feature pyramid network. *Scientific African*, 16:e01151, 2022.
[15] B Valarmathi, N Srinivasa Gupta, G Prakash, R Hemadri Reddy, S Saravanan, and P Shanmugasundaram. Hybrid deep learning algorithms for dog breed identification–a comparative analysis. *IEEE Access*, 2023.