Animal Species Recognition and Classification

Raktim Bijoypuri IIT2018125
Semester VI, Information Technology,
IIIT Allahabad, Prayagraj.

Abstract: This study aims to classify the images of various animal species using Machine Learning models. We did some research related to this and came up with some of our chosen research papers which we thought are worth studying before continuing our project. All these papers are one by one summarized below.

1. INTRODUCTION

Today, Machine Learning and deep Learning has so many applications. These models are made and then trained to make them achieve desired accuracy or to make them intelligent in a way. Artificial Intelligence has applications in almost every field. Image Classification is one of the Applications of machine learning. Image classification is a process in which we give an image input and get a class or category as output. This input is given to a machine learning model. There are many pre-trained models available such as VGG-16, ResNet50, Inceptionv3, EfficientNet, etc. Convolutional Neural Networks(CNN) remains the most preferred choice for image recognition and classification. We aim to develop an efficient image classifier which can predict the class i.e. species of the animals in an image.

2. BACKGROUND

Monitoring the animal species in their habitat can help us understand their migration patterns, habitat protection, etc. and can help ecologists and biologists in their studies and also in taking decisions regarding the same. We aim to make this thing as easy as we can. Since, taking photos manually and then processing them is a time consuming process. So, We aim to develop an efficient image classification system that can classify the images of various wild animal species. First we take a look at the related research papers previously available on the internet to understand the current technologies that are being used. This will help to finalise our methodology for the same.

3. Literature review

3.1 Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning[1] :

This paper uses a 2 step approach for recognising wild animals.
Using the SS dataset which consists of 3.2 million images of 48 species.
Step 1 :
Since camera trap images are used which respond to any kind of motion, therefore many of the images captured are actually blank, it was roughly

estimated that as much as 75% of the images are blank.
So in this step they attempt to prune the dataset by first recognising whether the image actually contains an animal or not.
This greatly reduces the size of the dataset.

For this task the paper studied and compared a variety of pre-trained models with transfer learning such as VGG16, ResNet etc.
For better results some basic pre-processing and data-augmentation techniques were applied.
In pre-processing first the images were scaled down to reduce the size of the dataset and to increase speed of the processing.
Further all the images were normalised by subtracting mean, and dividing by standard deviation of the pixel values.
In the data-augmentation part we try to make the model more robust by changing small features of the images to emulate real changes in conditions as much as possible.

Table S.6. The accuracy on capture events (as opposed to individual images) of models for task I: Detecting Images That Contain Animals.

Architecture	Top-1 accuracy for capture events
AlexNet	96.3%
NiN	96.6%
VGG	96.8%
GoogLeNet	96.9%
ResNet-18	96.8%
ResNet-34	96.8%
ResNet-50	97.1%
ResNet-101	96.8%
ResNet-152	96.8%

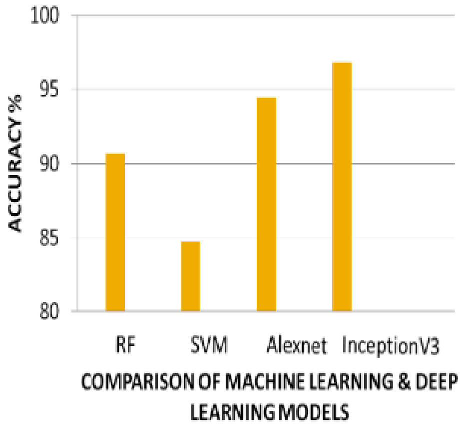
Step 2 -
Recognising which species the animal belongs to -
Each image was mapped to one of the possible 48 species, for this again several CNN's were used(such as VGG16,ResNet etc) and their results compared, for this task ResNet-152 gave the best results with 93.8% top-1 and 98.8% top-5.
The paper is also considering the top-5 accuracy because in cases where human intervention is required for labeling the images then top-5 can narrow the options for humans a lot, in fact the best top-5 accuracy obtained was 99.1%.

Table 2. Accuracy of different models on task I: Detecting images that contain animals

Architecture	Top-1 accuracy, %
AlexNet	95.8
NiN	96.0
VGG	**96.8**
GoogLeNet	96.3
ResNet-18	96.3
ResNet-34	96.2
ResNet-50	96.3
ResNet-101	96.1
ResNet-152	96.1
Ensemble of models	96.6

3.2 Recognition Of Animal Species On Camera Trap Images Using Machine Learning And Deep Learning Models[2] :

This paper also uses several models and compares their results similar to above however the dataset used by this paper is very small as compared to the previous one.
The KTH dataset used only contains 1740 images with 12 classes of images.
This paper used 3 different methods i.e SVM, Neural network(Inception V3) and Random forest.
Again the Neural network performed the best with accuracy of 97%, Random first and SVM gave accuracy of 91% and 84% respectively.
This paper helps us to conclude that deep learning models are in fact better than other machine learning models.



3.3 Animal Classification System Based on Image Processing & Support Vector Machine[7]:

This paper handles both the data gathering part and classifying the images, since we will be using an already existing dataset so we directly look at the classification part.

Before classification the paper uses ACE media project for handling feature extraction, the features extracted belong to the broad classes such as Color, Texture, Shape etc.

2 color descriptors, 2 shape descriptors and 2 texture descriptors were used as inputs to the SVM.

The overall accuracy used was around 80% for the images collected of cats,dogs and tigers.

3.4 Animal Recognition and Identification with Deep Convolutional Neural Networks for Automated Wildlife Monitoring[3]

This paper uses the wildlife spotter dataset with 3 species Bird, rat, bandicoot.

Again the paper uses 2 steps as earlier for data pre-processing and deleting empty images with no animals.

For recognising images with no animals an accuracy of 96% was obtained since only 30% of the images in wildlife spotter dataset contained images so if needed humans annotators can also perform species classification quickly.

After this several pre-trained neural networks such as VGG16 and ResNet are used with transfer learning, the best accuracy obtained was with ResNet-50 of 84.39%.

Model	Accuracy (%)	
	Training from scratch	*Fine-tuning*
Lite AlexNet	87.80	-
VGG-16	88.03	88.23
ResNet-50	87.97	76.43

3.5 Multiple feature detection of Predator animals using SVM and MLV classifiers:

This paper uses feature extraction of predator animals mainly from their face i.e. their eyes, ears, etc. because face contains most of the valuable information that is needed to classify the animals in two categories i.e. **Pets and Predators.**

The training set contained 150 images (75 images of predator animals and 75 images of pet animals). The testing set contained 50 images (25 images each of the predator class and the pet class).

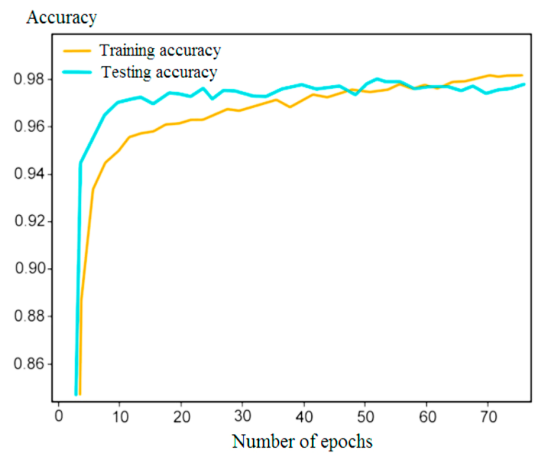
In this experiment, the results of both the SVM and the MLV classifier are computed. It is found that SVM average accuracy for each class is 78% and that of the MLV classifier is 82%.

3.6 Animal Species recognition in the wildlife based on muzzle and shape features using joint CNN:

In this paper, two models are applied on the same dataset and the results are compared. This paper uses a dataset captured by camera traps in Ergaki national park. The details of dataset are shown :

Species	Samples	Species	Samples
Red Fox	3123	Moose	959
Brown Bear	3732	Wild Boar	4332
Western Capercaillie	848	Badger	639
Lynx	1007	White Hare	3824
Maral	3510	Squirrel	118
Roe Deer	4228	Dog	512

First experiment was conducted using a single branch VGG19 and the second experiment was conducted using Joint CNN. Joint CNN out-performed VGG19 and attained better accuracy. The plot of accuracy of training and testing dataset of Joint CNN is shown :

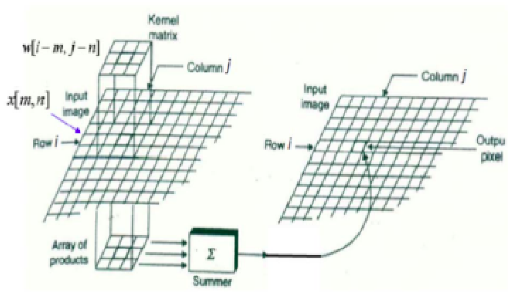


3.7 Varied channels region proposal and classification network for wildlife image classification under complex environment:

In this paper, the architecture of CNN is a varied channels region proposal and classification network. Speciality of this paper lies in the architecture of the model. The input image is fed to the model as different component in different channels. This helps to evaluate different parts separately resulting a great improvement in final accuracy of prediction. This model takes into account the movement of animal by their appearances and focuses mainly on the local features.

3.8 Understanding of a Convolutional Neural Network

This paper focuses on the method of how we can extract features from images of wild animals. The architecture used here is Convolutional Neural Network. In actual practice this is a deep neural network with multiple layers. The layers cummatively are called as hidden layers. Each layers has its own functionality. There are Convolutional layer, Maxpooling, Fully Connected layers etc. A non linearity layer is used to cut off the generated output. ReLu Functions are used in this case.



3.9 Supervised and Unsupervised Learning in Animal Classification

Dataset of 2000 animal images of 20 species was used, and both supervised and unsupervised methods were used for testing and finding the best model.
First dimensionality reduction was done for both the methods. Then for supervised LDA was used and results fed into symbolic classifiers and for unsupervised PCA was used with K-means clustering.

Training Percentage	With dimensionality reduction		Without dimensionality reduction	
	Supervised	*Unsupervised*	*Supervised*	*Unsupervised*
30	59.76	54.89	46.32	44.76
50	75.05	66.65	53.12	51.43
70	79.54	75.46	61.63	57.87

3.10 Machine learning to classify animal species in camera trap images: Applications in ecology

NCATI dataset of 3.7 million images was used for training a CNN (ResNet - 18) using tensorflow framework.
Top -1 accuracy of 97.6% was achieved with an accuracy of more than 99% for top-5.
For Model validation only 5,900 images were used of 4 species.
Further the SS dataset was used for testing how well the model performs on images it has not seen before.

3.11 Fast Human-Animal Detection from Highly Cluttered Camera-Trap Images Using Joint Background Modeling and Deep Learning Classification

This paper attempts to classify the foreground objects in images into 3 categories i.e animal, human, background. First block wise segmentation is done on the image to extract foreground objects, The foreground objects are then fed into a DCNN .
Dataset of 30,000 images was used with 10,000 images per category.
The resulting DCNN was very fast with an accuracy of 83%.

3.12 Marine Animal Detection and Recognition with Advanced Deep Learning Models

This paper tries to classify different fish species using images.
First foreground detection is done to detect fishes, once images with fishes are detected and different fishes are detected as foreground objects within the same image, the results are fed into a pre trained neural network which is then fine tuned.
The best top-1 accuracy of classification was 84.5% with BN-Inception model.
For foreground object detection of fishes ResNet-10 gave accuracy 71%.

3.13 Marine Animal Classification With Correntropy-Loss-Based Multi View Learning

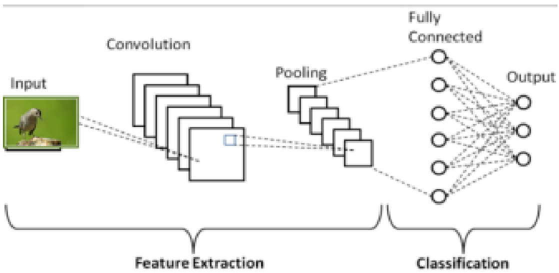
20,000 images of 23 species of fish was used for training, for this correntropy-loss-based multi-view learning was used.
Some features used were hand made whereas others were extracted.
The best accuracy of 95.4% was achieved for C-MV.

3.14 BIRD SPECIES CLASSIFICATION USING DEEP LEARNING APPROACH

In this paper, the authors have acquired images of 20 different bird species from the internet.

DATASET OF BIRD SPECIES WITH THE NUMBER COUNT		
SPECIES NAME	TRAIN IMAGES	TEST IMAGES
Black_footed_Albatross	345	87
Bobolink	320	80
Brewer_Blackbird	332	84
Cardinal	307	77
Crested_Aukle	336	84
Gray_Catbird	553	139
Groove_billed_Ani	336	84
Indigo_Bunting	313	79
Laysan_Albatross	332	84
Lazuli_Bunting	358	90
Least_Auklet	536	135
Painted_Bunting	369	93
Parakeet_Auklet	369	93
Red_winged_Blackbird	388	98
Rhinoceros_Auklet	356	90
Rusty_Blackbird	474	119
Sooty_Albatross	374	94
Spotted_Catbird	320	80
Yellow_breasted_Chat	336	84
Yellow_headed_Blackbird	313	79
Total	7967	1853

A deep convolutional network has been developed to classify these images.



An accurate classification of 98.75% is achieved with this model.

3.15 BIRD SPECIES CLASSIFICATION USING DEEP LEARNING APPROACH

In this paper, the classification has been done using unsupervised learning methods as well as deep learning CNN. Caltech-UCSD Birds 200 [CUB-200-2011] Dataset is used as a dataset. Various types of a;gorithm like - Pose Norm, Part-based R-CNN, Multiple granularity CNN, Diversified visual attention network (DVAN) is used. Accuracy obtained --

Sr. No	Method	Accuracy
1)	Pose Norm	80%
2)	Part-based R-CNN	76.4%
3)	Multiple granularity CNN	80%
4)	Diversified visual attention network (DVAN)	79%
5)	The deep LAC localization, alignment, and classification	80.3%
6)	Deep Learning Using CNN	85%

3.16 BIRD SPECIES CLASSIFICATION USING DEEP LEARNING APPROACH

In this paper CNN is used for feature extraction. Then these feature maps are fed into a SVM binary classifier. To fine tune the result stochastic gradient descent algorithm is used. The dataset used here is a manually captured 9312 image set. Data augmentation is used.

3.17 Pig Face Recognition using CNN:

In this paper, CNN model was trained using a dataset which consisted of 1553 images of 10 different pigs. The images were of face only so that it could identify each pig from its face. A dissimilar dataset was then tested by using 3 approaches, Fisherface(pre-trained), VGG-face(pre-trained), and their own CNN model. The results of all the three are shown below:

Method	Accuracy(%)	False Positive(%)	False Negative(%)
Fisherface	78.4	26.3	29.7
VGG-Face	91.0	10.3	13.9

CNN	96.7	5.8	5.7

3.18 A Block Based Approach for Animal classification using K-nearest neighbours(KNN) and Probabilistic neural networks(PNN):

In this paper, a method for animal classification is proposed. Firstly, the unwanted backgrounds were removed by segmentation. After segmentation, blocks of images are formed and then they have done feature extraction on the blocks. They conducted a test on their own dataset which consisted of 4000 images of animals of 25 different classes. The Accuracy results when 70% of the dataset is used as training are as shown:

Blocks	PNN(%)	KNN(%)
1	70.70	82.38
4	70.20	72.55
16	71.97	74.96
64	56.70	65.32

It can be noticed that KNN performed better than PNN.

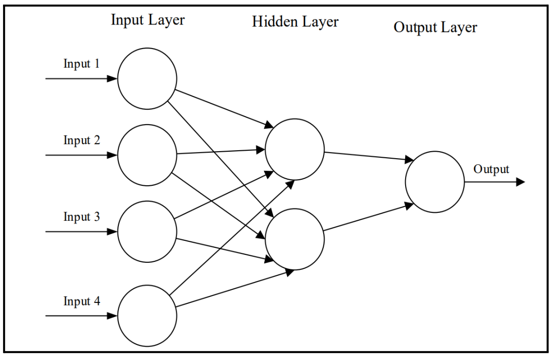
3.19 Identification of Animal species in camera trap images using Deep CNN:

Snapshot serengeti Dataset is used in this paper. This dataset contains the images of 48 different animal species and also it is a highly unbalanced dataset. Unbalanced dataset means that it contains more images of one species whereas contains very less images of other species. The results on this dataset are shown here :

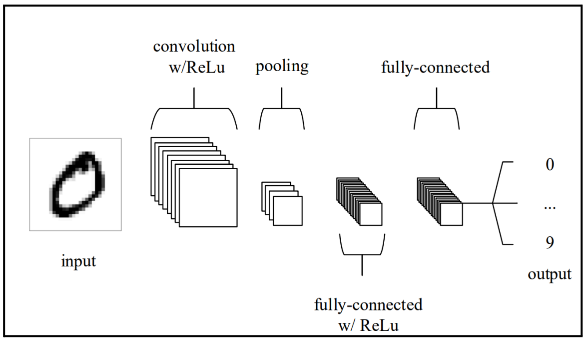
Specie	Accuracy[%]	Specie	Accuracy[%]
Baboon	98.3	Jackal	92.4
Buffalo	83.3	Kori bustard	97.9
Cheetah	97.0	Lion female&cub	83.3
Dik dik	75.8	Lion Male	77.0
Eland	87.5	Ostrich	94.1
Elephant	90.4	Reedbuck	95.0
Giraffe	96.6	Secretary bird	95.4
Grant's Gazelle	65.0	Spotted Hyena	85.4
Guinea fowl	99.5	Thomson's Gazelle	71.6
Hartebeest	97.5	Topi	95.8
Hippopotamus	94.1	Warthog	98.0
Human	99.1	Wildebeest	93.1
Impala	92.0	Zebra	99.5

3.20 Introduction to Neural Networks:

From this paper, we learnt how a basic Artificial Neural Network works. Their computational processing systems are heavily inspired by the functioning of a biological nervous system. The basis structure of a number of common ANNs is:



We read about supervised learning, unsupervised learning, overfitting, CNN, and its various layers. Basic CNN structure with its layers is shown below:



4. METHODOLOGY

Our methodology contains a few separate parts. We majorly classify them as 1) Image Preprocessing, 2) Data preprocessing and augmentation, 3) Model training and Prediction.

In this following we will be describing our methodology,

1) **Image Preprocessing:** The dataset we will be using images of animals at a high resolution and various sizes. The width and height varies from image to image. This is inconvenient for a machine learning model to learn the underlying pattern or extract the key features of this kind of dataset. To avoid this we will resize the image into a constant size. The multicolor images will be converted into grayscale to reduce the dimension of input data. We will crop the center part of the image to reduce as much background data as possible and only focus on the animal species.

Fig. Original

Fig. After image preprocessing

2) **Data Preprocessing and Augmentation:**

Animal Detection:
The aforementioned dataset contains images which are often of no use, i.e, there is no animal species on those images. Since the total dataset is very large to reduce the need of computation power and running time of the algorithm we will discard those images which contain no animals. As in reference to other papers 75% of the total dataset contains no animals. So this task will be done as a binary classification problem. We will call it an **animal detector.** In this task we will train a Convolutional Neural Network (hereafter referred as CNN) to classify the images into 2 sets. One that contains animals and one that does not contain animals.

Data augmentation: By filtering out the empty photos we substantially reduce the dataset. But to make a good prediction the model needs a greater amount of data. So for producing legitimate data for our model we will use **data augmentation techniques.** In actual practice the original images will be copied from different sides. By applying **flipping, rotating, shearing, cropping, zooming in, zooming out** and **changing contrast and brightness** etc. operations we will generate more legitimate data. This will help the model to identify the species of the animal more accurately.

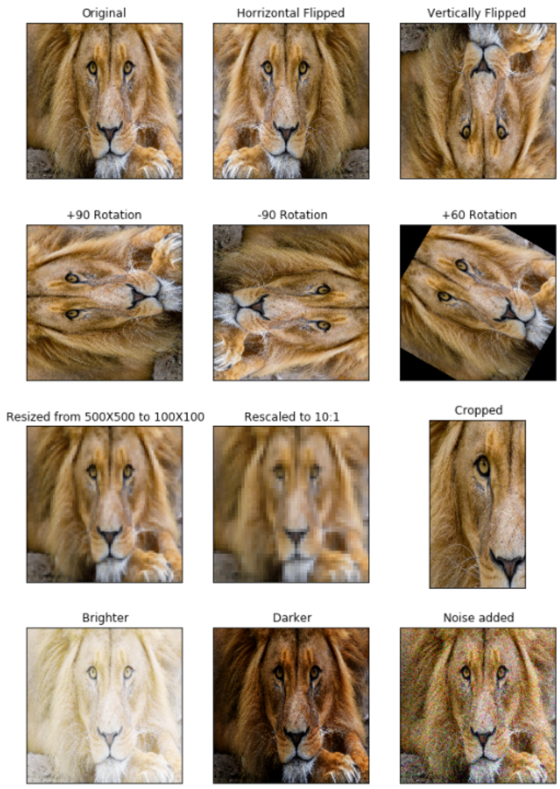


Fig. Data Augmentation example

3) Model training and Prediction:
This is the phase where we build our machine learning model and train and test upon our preprocessed dataset.

The dataset will be divided into 2 parts, namely train and test set. The amount of the split of the original set is subject to experimental results, normally we are thinking of doing a 80-20 split.

We will be using CNN to identify/classify the animal species. The accuracy of this model will depend on the underlying architecture of the CNN model. The number of convolutional layers, activation function, Max Pooling layers are subject to experimental results.

We can also use **Transfer learning**. It is a bleeding edge deep learning technique to train models on a small dataset. In this case we will use some **state of the art** pre-built CNN architectures like **VGG16, GoogleNet, ResNet50, ResNet101, AlexNet** etc. These models are pretrained on some other dataset. For example **VGG16** is trained on **ImageNet** which is one of the largest image dataset publicly available. We will use the parameters of these models for fine tuning of our model. This will help since we have a small dataset.

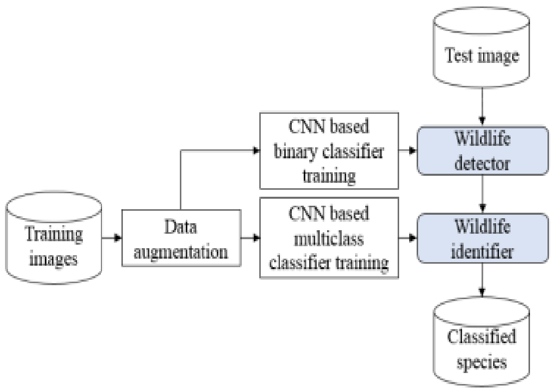


Fig. Diagram of Model

5. EXPERIMENTS AND RESULTS

We have used the Animals-10 dataset publicly available on kaggle. The dataset contains 10 different classes of animals. Each different class has 2000 - 5000 pictures. The total number of pictures in that dataset is around 29,000. This photos were taken from google.

<u>**Building model from scratch;**</u>

Initially, we build our own model from scratch. We implemented VGG16 architecture. We trained it on our Img1000 and Img5000 dataset. Tested in 3 validation set. But it gave very poor results.

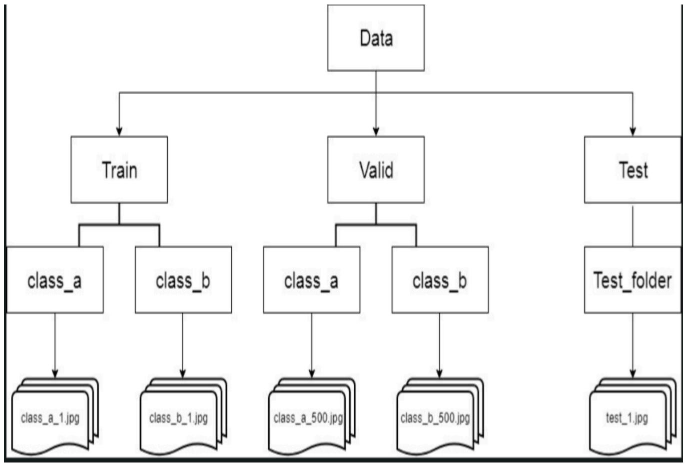
Model	Dataset	Avg accuracy
vgg16 from scratch	Img1000	24.5%
vgg16 from scratch	Img5000	48.67%

Since the number of pictures in each category is less, we have implemented Transfer Learning. In transfer learning a state of the art pre-trained model is used to predict the label of other datasets. Since the pre-trained model is already trained over huge datasets a small amount of data is enough to provide good accuracy.

In our case we have used VGG16 and ResNet50 pre-trained models. These models are trained on the ImageNet dataset. This dataset is a huge one with approximately 1.2 million different kind of photos. There are 1000 different classes.

Process:-

First we created 2 smaller dataset from the original one. In each dataset there are 3 folders train, validation, test. The directory structure looks like below-



dataset1 contains 1000 images for training, 200 images for validation and 100 images for testing.
dataset2 contains 5000 images for training, 1000 images for validation and 500 images for testing.

Step1:- We have used the Keras library for importing the pretrained models and their weights.

Step2:- The last/topmost layer of the model is removed and replaced with a Dense layer with 10 different classes and softmax activation.

Step3:- We kept all but the newly added layer freezed. These layers will not change their weights in the training process.

Step4:- After this we start the training process with certain batchsize, epochs and steps per epochs.

Step5:- After training we save the model in h5 format for future use.

Model Architecture:

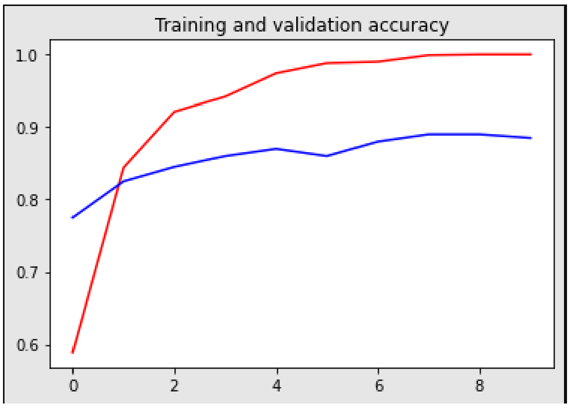
To get started with this process we create a keras sequential model which contains 2 layers. First layer is the pretrained model which is kept non-trainable and the second layer is a Dense(Fully-Connected) layer with softmax activation. Pretrained layer is kept non-trainable because we want to use the weights of that model. Dense layer contains 10 different classes as per our dataset.

Results:

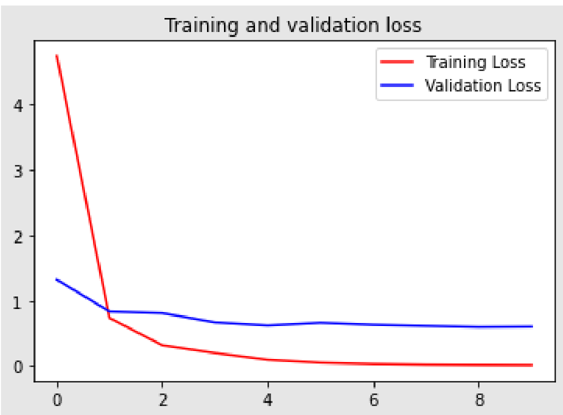
3 separate validation set are prepared to evaluate the models.

VGG16:

This model has given 99.83% accuracy on the training data, 88.50% accuracy on the validation data.



Training and validation accuracy vs epochs



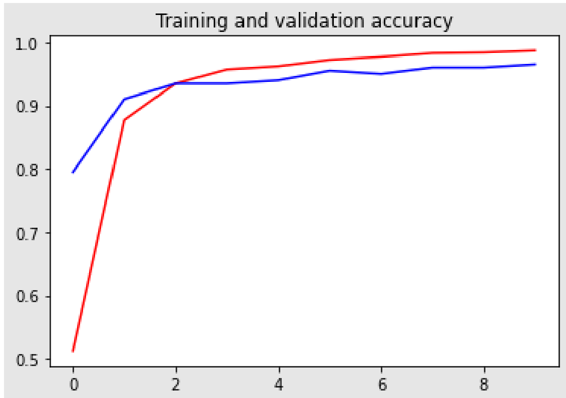
Training and validation loss vs epochs

We have evaluated this model on 3 validation sets which are not involved in training which gave us the following results.

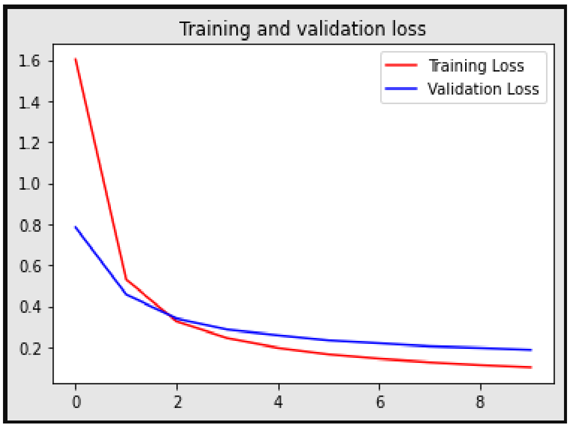
Validation Set	Accuracy
validation set1	81.99%
validation set2	83.49%
validation set3	80.5%

ResNet50:

This model has given 98.39% accuracy on the training data, 96.50% accuracy on the validation data.



Training and validation accuracy vs epochs



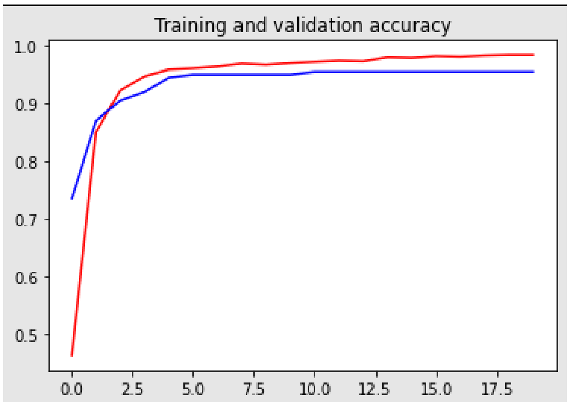
Training and validation loss vs epochs

We have evaluated this model on 3 validation sets which are not involved in training which gave us the following results.

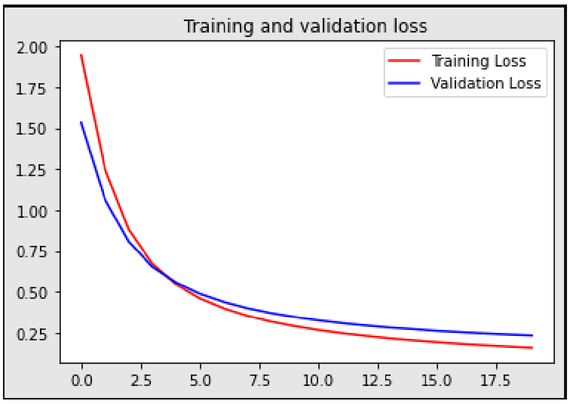
Validation Set	Accuracy
validation set1	88.99%
validation set2	89.49%
validation set3	86.5%

InceptionV3:

This model has given 98.97% accuracy on the training data, 95.50% accuracy on the validation data.



Training and validation accuracy vs epochs



Training and validation loss vs epochs

We have evaluated this model on 3 validation sets which are not involved in training which gave us the following results.

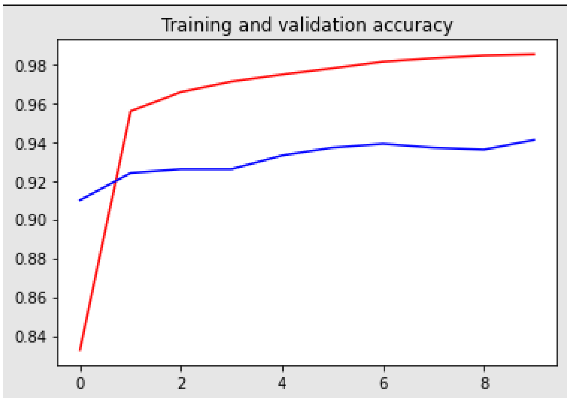
Validation Set	Accuracy
validation set1	94%
validation set2	93%
validation set3	93.5%

All of these results are obtained by using 100 images for each class with a total 1000 images for training and 200 images for validation (20 images per class).

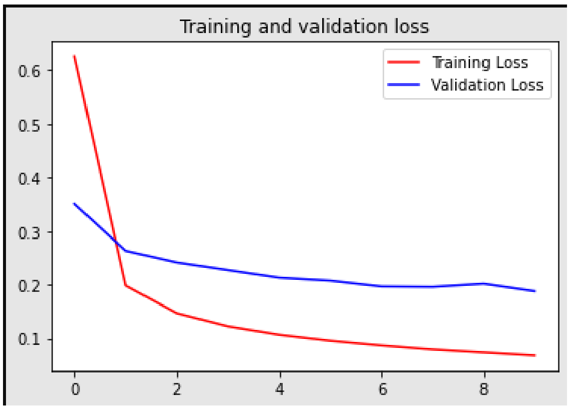
To improve the models we have decided to put more data into training. For this we select ResNet50 and InceptionV3 to work with. We gained the following results by using 5000 images in training and 1000 images for validation.

ResNet50:

This model has given 98.84% accuracy on the training data, 93.87% accuracy on the validation data.



Training and validation accuracy vs epochs



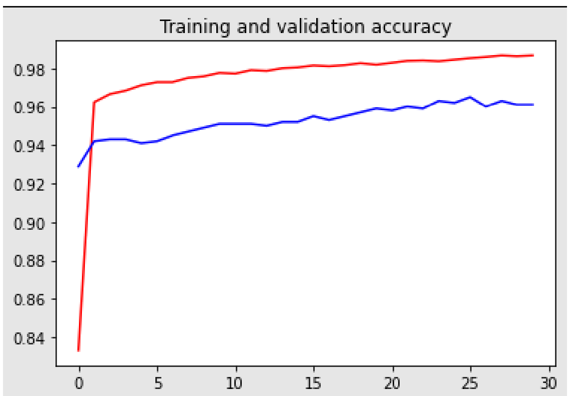
Training and validation loss vs epochs

We have evaluated this model on 3 validation sets which are not involved in training which gave us the following results.

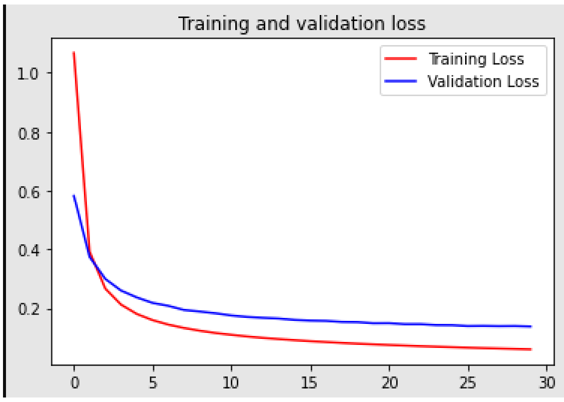
Validation Set	Accuracy
validation set1	97.12%
validation set2	95.01%
validation set3	93.97%

<u>InceptionV3:</u>

This model has given 98.27% accuracy on the training data, 96.20% accuracy on the validation data.



Training and validation accuracy vs epochs



Training and validation loss vs epochs

We have evaluated this model on 3 validation sets which are not involved in training which gave us the following results.

Validation Set	Accuracy
validation set1	98.9%
validation set2	96.12%
validation set3	95.34%

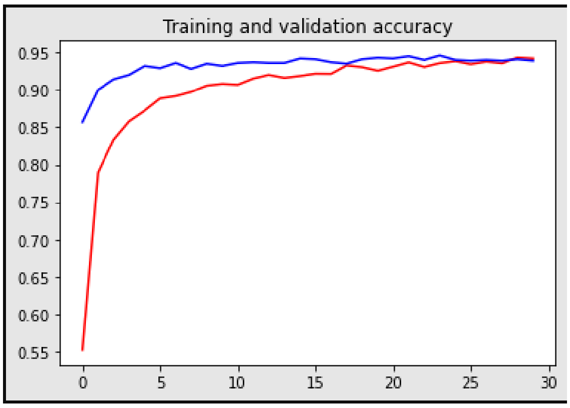
In our next phase of experiments we decide to fine tune our previously trained models. Because in our previous model training we achieve greater accuracy in model training than model validation, which indicates overfitting of the model.

<u>Fine Tuning of the models:</u>

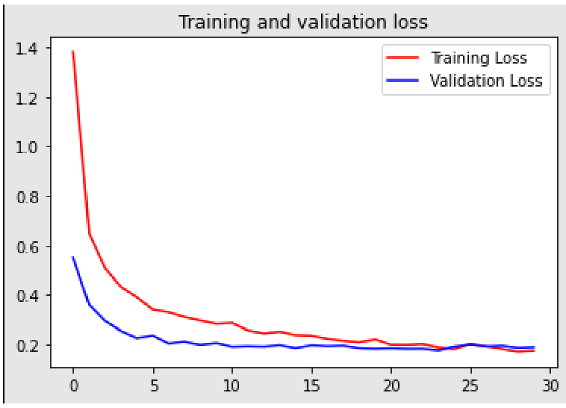
To fine tune our model and reduce the overfitting we make changes in our model architecture. Previously our architecture only consisted of an extra Dense layer over the top of the pretrained model. Now, we add some more layers in between them. Keras provides us with different layers like - Dropout, Dense, Flatten, MaxPooling etc. In this case the number and type of layers vary from model to model. We add as we see fit.

ResNet50:

This model has given 95.84% accuracy on the training data, 94.50% accuracy on the validation data.



Training and validation accuracy vs epochs



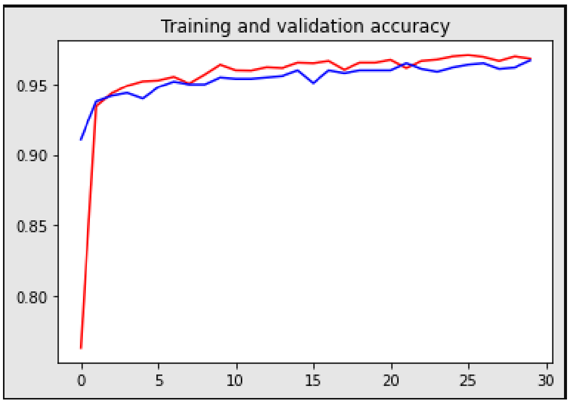
Training and validation loss vs epochs

We have evaluated this model on 3 validation sets which are not involved in training which gave us the following results.

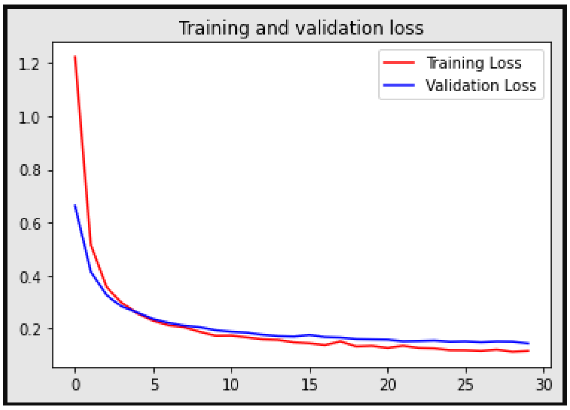
Validation Set	Accuracy
validation set1	95.98%
validation set2	95.5%
validation set3	93%

InceptionV3:

This model has given 97.84% accuracy on the training data, 96.90% accuracy on the validation data.



Training and validation accuracy vs epochs



Training and validation loss vs epochs

We have evaluated this model on 3 validation sets which are not involved in training which gave us the following results.

Validation Set	Accuracy
validation set1	98.10%
validation set2	97.02%
validation set3	96.5%

We have tested all of these models on another dataset cats and dogs.

Model	Training Data	Accuracy
vgg16	Img1000	82.92%
ResNet50	Img1000	92.8%
ResNet50 + Fine Tuning	Img5000	96.09%
InceptionV3	Img1000	96.66%
InceptionV3 + FineTuning	Img5000	97.81%

5. CONCLUSION

Although we have implemented two of the best state of the art model publicly available, there are still some models that are still remain to be experimented by us like- InceptionV3, ResNet101 etc. Apart from that for future work we can fine tune these models, change the hyper parameters, include more data in training to get better accuracy.

6. REFERENCES

[1] Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning(2018).
Mohammad Sadegh Norouzzadeh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S. Palmer, Craig Packer, Jeff Clune

[2] Recognition Of Animal Species On Camera Trap Images Using Machine Learning And Deep Learning Models(2019).
Rajasekaran Thangarasu,Vishnu kumar Kaliappan ,Raguvaran Surendran,Kandasamy Sellamuthu ,Jayasheelan Palanisamy

[3] Animal Recognition and Identification with Deep Convolutional Neural Networks for Automated Wildlife Monitoring(2017)
Hung Nguyen, Sarah J. Maclagan, Tu Dinh Nguyen1, Thin Nguyen,
Paul Flemons, Kylie Andrews, Euan G. Ritchie Dinh Phung

[4] Alharbi, Fahad & Alharbi, Abrar & Kamioka, Eiji. (2019). Animal species classification using machine learning techniques. MATEC Web of Conferences. 277. 02033. 10.1051/matecconf/201927702033.

[5] Albawi, Saad & Abed Mohammed, Tareq & ALZAWI, Saad. (2017). Understanding of a Convolutional Neural Network. 10.1109/ICEngTechnol.2017.8308186.

[6] C. Bowley, A. Andes, S. Ellis-Felege and T. Desell, "Detecting wildlife in uncontrolled outdoor video using convolutional neural networks," 2016 IEEE 12th International Conference on e-Science (e-Science), Baltimore, MD, 2016, pp. 251-259, doi: 10.1109/eScience.2016.7870906.

[7] Animal Classification System Based on Image Processing & Support Vector Machine(2016)
A. W. D. Udaya Shalika, Lasantha Seneviratne.

[8] Supervised and Unsupervised Learning in Animal Classification(2016)

Manohar N , Sharath Kumar Y H , and Hemantha Kumar G

[9] Machine learning to classify animal species in camera trap images: Applications in ecology(2018)

[10]Fast Human-Animal Detection from Highly Cluttered Camera Trap Images Using Joint Background Modeling and Deep Learning Classification(2017).
Hayder Yousif, Jianhe Yuan Roland Kays Zhihai He

[11]Marine Animal Detection and Recognition with Advanced Deep Learning Models(2017)
Peiqin Zhuang, Linjie Xing, Yanlin Liu, Sheng Guo, Yu Qiao

[12]Marine Animal Classification With Correntropy-Loss-Based Multi View Learning(2019)
Zheng Cao , Shujian Yu , Bing Ouyang, Member, IEEE, Fraser Dalgleish , Anni Vuorenkoski, Gabriel Alsenas,
and Jose C. Principe

[13] Prof. Pralhad Gavali , Ms. Prachi Abhijeet Mhetre , Ms. Neha Chandrakhant Patil , Ms. Nikita Suresh Bamane, Ms. Harshal Dipak Buva, 2019, Bird Species Identification using Deep Learning, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 08, Issue 04 (April – 2019),

[14] SHRIHARSHA, TUSHARA, VIJETH, SURAJ, Dr. HEMAVATHI P 2020, BIRD SPECIES CLASSIFICATION USING DEEP LEARNING APPROACH INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT), Volume 07, Issue 04 (April – 2020)

[15] Niemi, J.; Tanttu, J.T. Deep Learning Case Study for Automatic Bird Identification. Appl. Sci. 2018, 8, 2089. https://doi.org/10.3390/app8112089

[16] Willi, Marco & Pitman, Ross & Cardoso, Anabelle & Locke, Christina & Swanson, Alexandra & Boyer, Amy & Veldthuis, Marten & Fortson, Lucy. (2018). Identifying Animal Species in Camera Trap Images using Deep Learning and Citizen Science. Methods in Ecology and Evolution. 10. 10.1111/2041-210x.13099

[17] Y.H. Sharath Kumar, N. Manohar, H.K. Chethan, Animal Classification System: A Block Based Approach, Procedia Computer Science, Volume 45, 2015, Pages 336-343, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2015.03.156.

[18] O'Shea, Keiron & Nash, Ryan. (2015). An Introduction to Convolutional Neural Networks. ArXiv e-prints.

[19] Margarita Favorskaya, Andrey Pakhirka, Animal species recognition in the wildlife based on muzzle and shape features using joint CNN, Procedia Computer Science, Volume 159, 2019, Pages 933-942, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2019.09.260.

[20] Guo, Yanhui & Rothfus, Thomas & Ashour, Amira S. & Si, Lei & Chunlai, du & Ting, Tih-Fen. (2019). A Varied Channels Region Proposal and Classification Network for Wildlife Image Classification Under Complex Environment. IET Image Processing. 14. 10.1049/iet-ipr.2019.1042.