Classification of Animal Images with modified CNN model

Nanda Kishor Jana[1], AbirChowdhury[2], DolaSaha[3], Chiranjib Dutta [4], Ananjan Maiti[5]

[1,2] Department of Computer Applications
Guru Nanak Institute of Technology
157/F Nilgunj Road, Panihati, Kolkata – 700114, West Bengal
3Assistant Professor, Dept. of Computer Applications
Guru Nanak Institute of Technology
157/F Nilgunj Road, Panihati, Kolkata – 700114, West Bengal

4Assistant Professor, Dept. of Computer Applications
Guru Nanak Institute of Technology
157/F Nilgunj Road, Panihati, Kolkata – 700114, West Bengal

5Assistant Professor, Dept. of Computer Science and Engineering
Guru Nanak Institute of Technology
157/F Nilgunj Road, Panihati, Kolkata – 700114, West Bengal

*Corresponding author's emails:*nanda.jana2018@gmail.com

Abstract- ***In the age of modern artificial intelligence, new methodologies are evolving day by day for species classification. There is some need for categorization of the animal for the preservation and restoration of species. Due to this, there are several techniques for the identification of animals out of the deep learning methods that are most useful for animal classification for their images. In this study, the modified and improved convolution neural network (CNN) has been employed. This study is the identification methodologies of the ten animals which are cat, dog, tiger, lion, sheep, rhino, cheetah, elephant, squirrel, and panda. These animals need to be identified by the artificial intelligence-based system for the large-scale preservation system and the accuracy obtained by the modified CNN is this. In the future, this study is going to evolve deep learning for human-less classification systems and this study will maintain the balance between the machine and animals in restricted areas that humans can't reach there. Animal classification is one of the core problems in Computer vision. A lot of attention has been associated with Deep Learning, specifically neural networks such as CNN. This animal classification model gives some accuracy of 95%.***

Keywords- Animal image classification, Anima image identification, Deep learning

1.0 Introduction

Monitoring wildlife is essential for determining how animals use their habitat. Different technologies have been developed to keep track on wild animals. For the purpose of informing choices on the management and conservation of animal species, effective and trustworthy. It is essential to monitor wild animals in their natural environments. Picture classification is the process of classifying and labelling groups of pixels or vectors in an image according to specified rules (Taheri,2017). CNN is a deep learning (DL) technique that takes an image as input, applies convolution, pooling, and normalizing to it, then outputs the processed picture. Because of their high accuracy, CNNs are used for image classification and recognition (Yousif,2017). In order to align our dataset with the design, transfer learning is used in this study, which is focuses on employing a “DEEP LEARNING” pre-trained architecture model. Mobile networks are one example of pre-trained designs that are employed. The data-set is initially separated into three separated group like train, test and validation. Using train and validation datasets, We test the dataset using to test dataset after training. The graph of accuracy and loss are created based on the training outcomes. Finally, based on the test dataset results, The confusion matrix can be used to measure accuracy. All architectures go through the same procedure. Each model's performance is contrasted in the findings section.

1.1 Previous Work

Deep learning is currently employed in a number of disciplines and research projects. The following are the most prominent Convolution neural nets (Convents) for image categorization and popularity. A significant development in computer vision is Alex Net. One of the first and most significant models that enforced 1x1 Conv's in order to generate additional functionality to the Conv-layers' possibilities was Network-in-Network, which was proposed in. The most popular deep learning architectures for image categorization include Google Net, Alex Net, Res Net, and Inception-1, 2, and 3. These architectures comprise 16 convolution layers, lots of pooled layers maximum, and three fully connected layers. The Google Net model was developed to deliver excellent accuracy while using little computational power. The Res Net's residual learning architecture has achieved successful results by connecting the ‘combined layer of convolution’ output and matching original input. By linking each layer to a distinct layer in a feed-forward fashion, Dense Conv-Network (Dense Net) has improved classification task performance. Animal differentiation was done in [10] using a variety of deep learning architectures, including Alex Net, Nin, and Reset. 'MODIFIED CNN' has provided good accuracy of 95.0 percent among the 6 architectures. In our study, animal species photos are classified using deep learning architectures such "MODIFIED CNN" Resnet, InceptionV3, and Mobile Net. Inception delivered the highest accuracy of the four architectures, at 95%, followed by mobile net, which came in second with 90%, "MODIFIED CNN," which came in third with 87 percent, and resent, which came in third with 53%. The 'MODIFIED CNN' design is the focus of this article; As a result, the same process is used for many architectures.

1.2 Experiment

Experiment Dataset **-,** this study is to evaluate our**,** this study does pixel mean processing to the image, which not only removes the average intensity of the image but also removes common parts.

2.0 Methodology

The study of animal species image classification (Tiwari, 2020) has previously been done by many researchers (Prudhivi, 2021). With the images**,** this study tried to classify them with the CNN process and with the reference datasets**,** this study does pixels mean processing to the images.

2.1 Data Augmentation–

This work suggests a brand-new offline data augmentation technique that turns an input image into a new, single image with various copies placed in n n cells and randomly rotated. When n is used with a higher number, a new picture with (Okafor, 2017) more diverse stances results. Because utilizing greater numbers of n caused the cow pictures to seem to be very tiny, the value of n was set to 4 in the studies. Figure provides an example of the suggested data-augmentation technique and the overall classification system utilizing CNN. The many adjustments made to the source image to create the multi-orientation image are described by the pseudo-code in Algorithm 1. Background pixels are added to the newly generated image after the images have been inserted to bind them together. The edges of the photos are used for this by employing nearest-neighbor pixels (Mikoajczyk, 2018).

2.2 Model Architecture –

In deep learning, there are numerous pre-trained architectures (Willi, 2017). In this study, the first sequential model is built using one of them, the "MODIFIED CNN" design. The model is then synchronized to the "MODIFIED CNN" design. It is among the most superior computer vision model designs ever created. Instead of using a large range of hyper-parameters, the "MODIFIED CNN" employed a 3x3 kernel with stride one convolution layers and consistently used the same padding and maypole layer of a pair of 2x2 stride two kernels. Throughout the whole architecture, the convolution and pooling layers are deliberately placed. After the pair of fully connected layers, a soft-max function for the output is presented. Fig.2 shows the suggested architecture in action. A kernel in the Convolution 2D layer is a little matrix. It is used to carry out tasks including edge detection, sharpening, and blurring, among others. Convolution between a kernel and the query picture is used to accomplish this. Features that were retrieved from the previous layer make up the output. The number of channels increases with each convolution layer.

CNN Model description

A distinct CNN model with a batch size of 10 and 5 training epochs was proposed in this study. In this study, the image size is 512 x 512. Ten classes have been used to train the model. For classifying, the categorical cross-entropy function was helpful. Loss error is minimized using the Adam optimizer. Before each training, random shuffling and image preparation, such as flipping and normalization, were organized in order to assist the model to be tuned more effectively. The formal architecture of the suggested model for the classification of the novel animal dataset used in this study is depicted in the following figure 1.

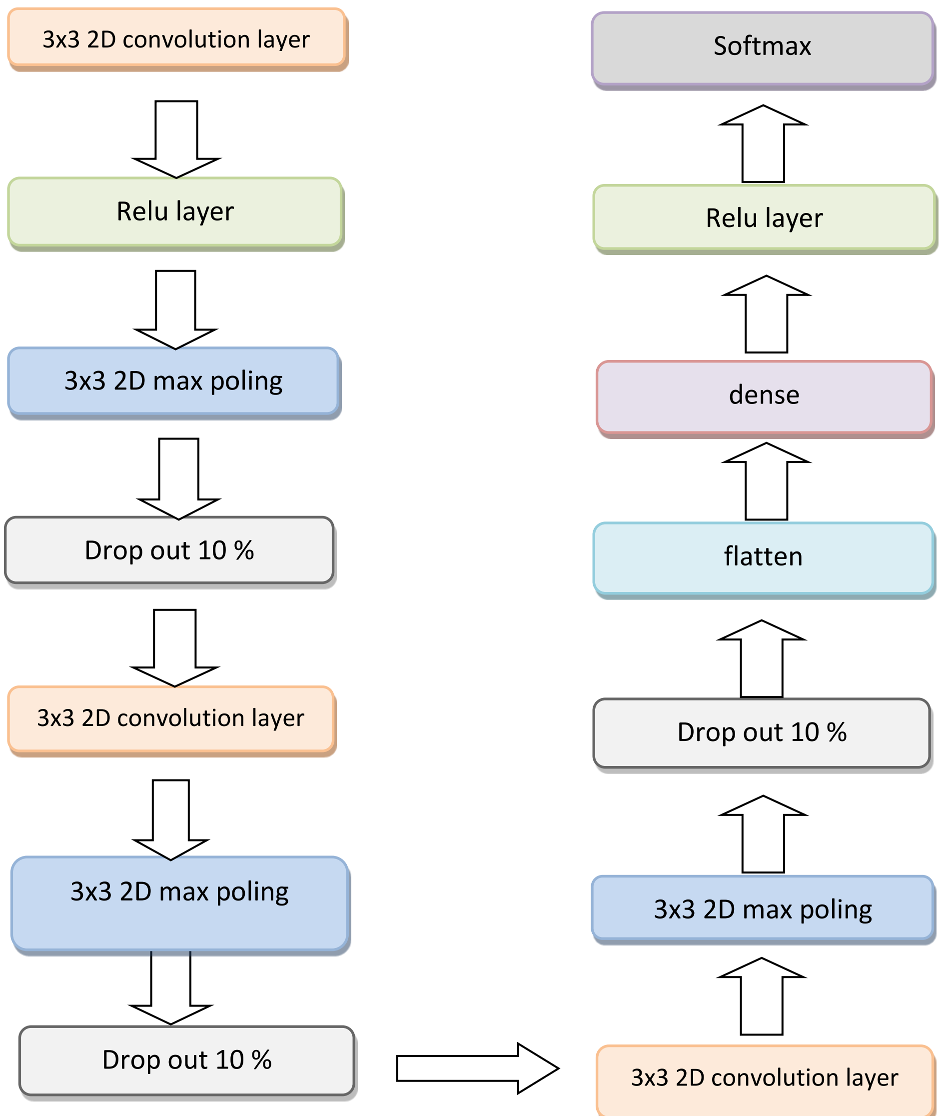

Fig.1 The Architecture of the Proposed CNN Model

3.0 Result Analysis

A. Accuracy **B. Loss**

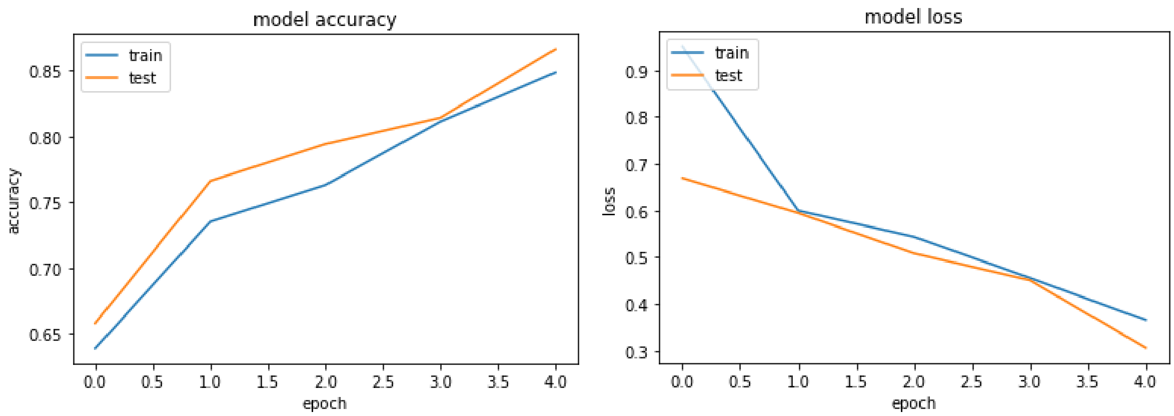

Fig 2. (a) accuracy of the model and (b) loss of the model

Fig. 2 training and testing detail:

When we done all processes then it gives us some graph, it called to be accuracy graph. Here we see two lines as per images and here two lines are representing two names i. train, ii. Test. Orange line for value accuracy and blue line for value loss. When the orange line increase then it represents the highly accuracy, when the orange line decreases then it represents the shortly loss. At the same process when the blue line increase then it represents the highly loss, when the blue line decreases then it represents the shortly loss.

A. Loss **B. Accuracy**

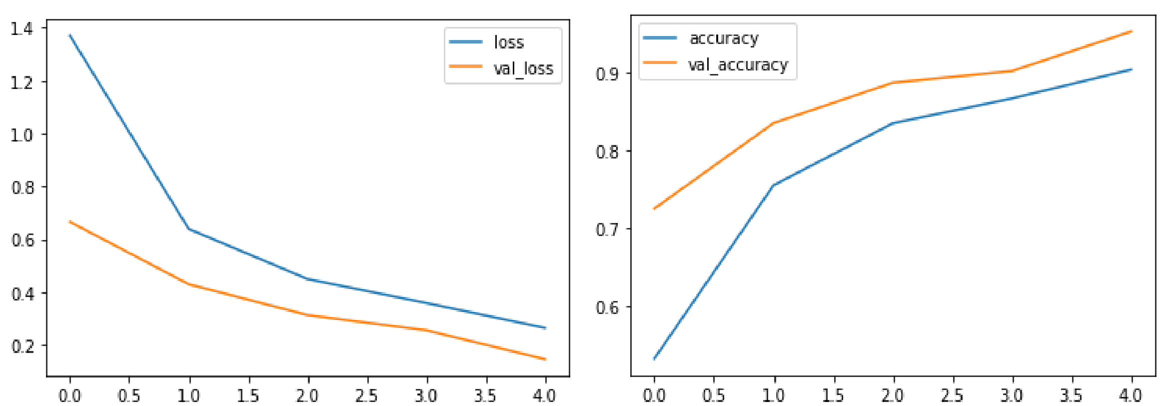

Fig 3. (a)loss of the model and (b) accuracy of the model

Fig. 3 Loss and Accuracy

Here in this graph two images are mainly representing two lines on are orange and one are blue here two lines represents are two particular names i. accuracy ii. Loss. When the blue line increase then it represents the highly accuracy, when the blue line decreases then it represents the shortly loss. When the orange line increase then it represents the highly accuracy, when the orange line decreases then it represents the shortly accuracy.

3.1 graphs of accuracy and loss

At the intersection of the train and validate graphs, the training accuracy continued to increase while the validation accuracy stabilized. Fig. 1. At the intersection of the training and validation graphs, the training loss continued to decrease while the validation loss became constant.

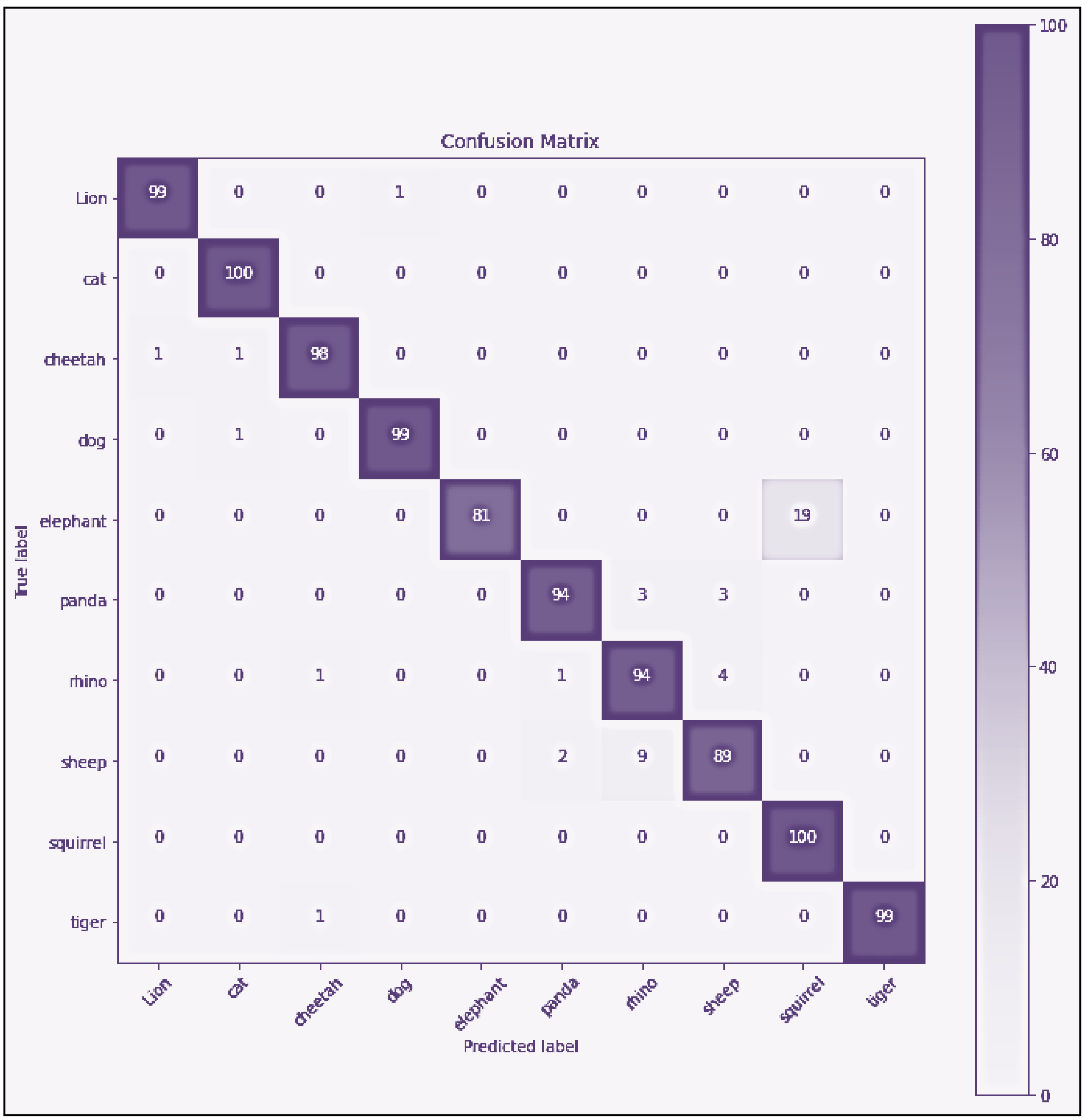

Fig:3 Confusion matrix of 'MODIFIED CNN'

Here we see in this table shown us the names of animals (Lion, Cat, Cheetah, Dog, Elephant, Panda, Rhino, Sheep, Squirrel, Tiger) Fig. 3 is represents us the confusion matrix of all species animals which is shown in this table. Here we see the all-accuracy level of this species. Fig.3 which is the main figure of CNN model, it's called the confusion matrix.

Table 1:

Val loss:	0.14268039166927338
Val accuracy:	0.9520000219345093
Test loss:	0.14685308933258057
Test accuracy:	0.9470000267028809

represent the value accuracy, value loss, test accuracy, test loss of ten animals Lion, Cat, Cheetah, Dog, Elephant, Panda, Rhino, Sheep, Squirrel, Tiger in this project. This is the configuration table of animals' classification CNN model. We always try to gather the whole pictures in a batch then we gave a picture of our choice. Then it finds that image which we gave in this batch. There after it gave us the value accuracy, value loss, test accuracy, test loss.

Table 2:

Animal Type	**Precision**	**recall**	**f1-score**	**support**
Lion	0.99	0.99	0.99	100
Cat	0.98	1.00	0.99	100
Cheetah	0.98	0.98	0.98	100
Dog	0.99	0.99	0.99	100
Elephant	1.00	0.81	0.90	100
Panda	0.97	0.94	0.95	100
Rhino	0.89	0.94	0.91	100
Sheep	0.93	0.89	0.91	100
Squirrel	0.84	1.00	0.91	100
Tiger	1.00	0.99	0.99	100

Table 2: Summary of the model without Pre trained Architectures

first of all, we take ten species of animals 1000 pictures of per animal. These animals are Lion, Cat, Cheetah, Dog, Elephant, Panda, Rhino, Sheep, Squirrel, Tiger. We separate 10 % pictures within 100% (100 out of 1000). Table 2. Represents the percentage table of animal types.

Table 3:

Accuracy	0.95	1000		
Minimum avg	0.96	0.95	0.95	1000
Maximum Avg	0.96	0.95	0.95	1000

Represents the whole accuracy of animals it had two parts i. Maximum average ii. Minimum average. Maximum average also represents the maximum accuracy of animals' species photos, Minimum average also represent the minimum accuracy of animals' species photos

3.3 Experimental environment

The experiment was carried out on the operating system with the system version 18.04, and the graphics card was GeForce RTX 1650. The deep learning framework used in the experiment is Tensor flow, and the language is Python.

3.4 Dataset description-

The photographs of the animals in the dataset were downloaded from the internet. Images of many animals, including CAT, DOG, LION, CHEETAH, TIGER, SQUIRREL, RHINO, ELEPHANT, SHEEP, and PANDA, are included in the dataset. Ten thousand photos make up the dataset, which is broken up into train, validation, and evolution datasets. There are 1000 photos of each animal or 1000x10 photographs. Following that, 1000x10 images are classified as training images, 100x10 images are validation images, and 100x10 images are demonstrated as evolution images. After multiple models were used to train the model on the dataset, InceptionV3 had the highest accuracy.

3.5 matrix of confusion

The confusion matrix shows the model's validation using test pictures. Determining the accuracy also benefits by visualizing the true positives, true negatives, false positives, and true negatives. Therefore, it is also possible to determine performance. The "MODIFIED CNN" architecture provided the confusion matrix. In Fig. 3, the outcomes of a sample of photos are displayed. More than 90% of the sample photos were correctly classified. Fig. 3 displays several examples of the model's outputs.

3.6 Comparative analysis

This study also used convolution layers and layer pooling to carry out the job without transferring learning. Fig. 3 depicts the model's executive summary. This model's accuracy has dropped compared to the model built using deep learning after being trained on the same dataset. Fig. 7(a) and (b) show the accuracy and loss graphs, respectively (b). When compared to Inceptionv3, which had the best accuracy of all pre-trained networks at 95%, this model's accuracy of 30% is considerably lower.

3.7 Conclusion -

The proposed framework works for the different types of animals with the image of the 512 x 512 image set. There are some researches available for this domain for the identification of different images of different animals. This study focused on different types of animals for remote monitoring and identification of their activities. The classification result shows good accuracy for different types of species as cat, dog, sheep, tiger, lion, rhino, elephant, panda, squirrel, and cheetah are dangerous. It has formed 85 to 95% accuracy in the future there could be different challenges for different environmental species for future development of this particular deep learning model.

References

Mikołajczyk, A., &Grochowski, M. (2018, May). Data augmentation for improving deep learning in image classification problem. In *2018 international interdisciplinary PhD workshop (IIPhDW)* (pp. 117-122). IEEE.

Okafor, E., Smit, R., Schomaker, L., &Wiering, M. (2017, July). Operational data augmentation in classifying single aerial images of animals. In *2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA)* (pp. 354-360). IEEE.

Prudhivi, L., Narayana, M., Subrahmanyam, C., & Krishna, M. G. (2021). Animal species image classification. *Materials Today: Proceedings*.

Tiwari, V., Pandey, C., Dwivedi, A., & Yadav, V. (2020, December). Image classification using deep neural network. In *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)* (pp. 730-733). IEEE.

Taheri, S., &Toygar, Ö. (2018). Animal classification using facial images with score-level fusion. *IET computer vision*, *12*(5), 679-685.

Willi, M., Pitman, R. T., Cardoso, A. W., Locke, C., Swanson, A., Boyer, A., ... & Fortson, L. (2019). Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution*, *10*(1), 80-91.

Yousif, H., Yuan, J., Kays, R., & He, Z. (2017, May). Fast human-animal detection from highly cluttered camera-trap images using joint background modeling and deep learning classification. In *2017 IEEE international symposium on circuits and systems (ISCAS)* (pp. 1-4). IEEE.