J. Inst. Eng. India Ser. B
https://doi.org/10.1007/s40031-020-00471-8

ORIGINAL CONTRIBUTION

InceptGI: a ConvNet-Based Classification Model for Identifying Goat Breeds in India

Satyendra Nath Mandal[1] · Pritam Ghosh[1] · Kaushik Mukherjee[1] · Sanket Dan[1] · Subhranil Mustafi[1] · Kunal Roy[1] · Dilip Kumar Hajra[2] · Santanu Banik[3]

Received: 12 May 2020 / Accepted: 11 July 2020
© The Institution of Engineers (India) 2020

Abstract In this paper, an attempt has been made to develop a model to decide with precision the breed identity of individual goat by using its image. For image-based multi-class classification tasks, CNNs have been found to be the best tool. But selecting the most efficient CNN model for a particular classification scenario is a very difficult job. To find an optimal CNN model for goat breed prediction, we have compared two of the most popular pre-trained deep-learning-based CNN models (VGG-16 & Inception-v3) based on their performance. Both the models have been fine-tuned using transfer learning on the goat breed database. This goat breed database has been created from goat images of six different breeds, which have been captured from different organized registered goat farms in India and almost two thousand digital images of individual goat have been captured without imposing stress to animals. It has been observed that Inception-v3 has outperformed VGG-16 with higher accuracy and lower training time. To measure the prediction performance of this fine-tuned Inception-v3 model, it has been applied to a test set of pure breed goat images and standardized classification performance evaluation metrics have been used to evaluate the prediction results. From the results, it is established that the proposed method used in this paper is able to accurately classify (recognize) goat breeds with high accuracy. Finally, comparison has been made with prediction accuracies of different technologies used for identification of domestic animals.

Keywords Convolutional neural network · Goat image dataset · Data augmentation · Fine-tuning · Breed prediction and transfer learning

Introduction

Humans recognize objects by their visual cortex systems. The visual cortex of the brain, contained in the occipital lobe which is located at the back of the skull, is a part of the cerebral cortex that plays an important role in processing visual information. Visual information, as percepted, travels through the eyes, goes through a series of brain structures and reaches the visual cortex known as inferior temporal gyrus (IT) for object recognition. Inferior temporal gyrus (IT) is responsible for identifying the object based on the color and form of the object by comparing processed information for that object with the stored information in brain like face recognition [1]. In this way, people (farmers, experts and other stakeholders) can recognize goat breeds if they frequently observe individual animals of different goat breeds. But the accuracy of breed recognition in this way is dependent on many factors. More than one breed of goats may look alike. Differentiating similar looking breeds depends on the observation power of experts. The presence of variation within a breed in its breeding tract for many cases produces confusion even to the experts to declare the breed identity. Identity of breed of a goat is a specific demand by a farmer for various reasons which include marketing, breeding and linkages

✉ Satyendra Nath Mandal
satyen_kgec@rediffmail.com

1 Department of Information Technology, Kalyani Government Engineering College, Kalyani, Nadia, West Bengal 741235, India

2 Department of Agronomy, Faculty of Agriculture, UBKV, Pundibari, Cooch BeharWest Bengal, 736165, India

3 ICAR-National Research Centre on Pig, Rani, Guwahati, Assam 781131, India

Published online: 25 July 2020

with community-based programs. There is no ready to use technology for the purpose of declaring breed identity of goat for small-scale farming community who are the major contributor of total gene pool of the goat genetic resource in India.

Research evidence suggests that between 20% and 70%, averaging about 50%, of the genetic diversity within each species is distributed at the breed level which has been the result of 10,000 years of domestication. The genetic diversity within and between breed is a key to strategize food production in the years to come to fulfill the growing demand of ever-growing human population. With the aim, the FAO initiated a Special Action Program on Global Animal Genetic Resources in 1993 with participation of 160 member countries.

Declaration of breed identity by phenotypic characters many times seems unsatisfactory especially, when purity of the breed is concerned. Genotypic characterization involves availability of expensive laboratory facilities. As a result, a new method for breed recognition is needed which should be accurate, economically viable and can be easily availed by the common people. This was the motivation behind this research.

The aim of this study is to develop a model which stimulates human visual cortex system using convolutional neural network for deciding breed identity of individual goat from their images. The model consists of five modules which act as layered structure. The images of individual goats from six goat breeds named Blackbengal, Beetal, Jamunapari, Barbari, Jakhrana and Sirohi were captured using cell phone camera (Samsung A6). All the captured images were resized to get file with uniform property before storing into the Image database. The images were captured from organized farms and field with uniform (UB) or non-uniform (NUB) background.

In the database creation module, the pre-processed images of different breeds were randomly separated into training set, validation set and test Set. The train set images were augmented to get uniform data for all the breeds.

The selected combination of images was fed into model selection module for selection of the best CNN model with less complex and fairly good execution capability for identifying goat breed images. The train and validation images were sent to two CNN networks. The VGG-16 model and the Inception-v3 model, both trained in transfer learning fashion. Based on their performance (better accuracy and lower time to train), Inception-v3 was found to be the best model for goat breed prediction. Finally, the fully trained fine-tuned Inception-v3 model was applied to test set images from Goat Image database and its prediction performance has been evaluated. Compared to various other research done for animal breed prediction, this method proved to be the best.

The effort of classifying goat breeds by computer vision would help create global information system and would empower all the stakeholders including small farmers.

This paper is written in eleven sections. The first section is the introduction. In "Background Study" section, relevant background information is given. In "Image Capturing & Goat Image" section, the goat image capture process and database creation process are explained. In "Convolutional Neural Network" and "Transfer Learning" sections, the theory about CNN and transfer learning is illustrated. In "Performance Evaluation" section, the techniques used for evaluating the performance of the model is demonstrated and in "Proposed Model" section, the obtained results and its consequences are discussed. Lastly in "Conclusion & Future Work" section the concluding remarks are stated and in "Conclusion & Future Work" section the acknowledgements, followed by the references used in this paper.

Background Study

Hailu used morphological and molecular markers for nondescript breed identification and characterization, and also demonstrated the difficulty to use such methods in animal breed identification [2]. Peng et al. performed chicken breed identification through DNA barcoding which is based on molecular identification [3]. Mekonnen et al. performed cattle breed characterization through morphometric and qualitative traits [4]. Morphometric characterization of goat breed has been done by Jai et al. [5]. Kumar et al. emphasizes that new methods based on detection of visible features and phenotypic appearances of species is very important for study of animal biometric characteristics, study of animal trajectory and behavior analysis of species [6]. This led to use of neural networks for animal breed identification since they are able to extract visible features from animals very efficiently. Krizhevsky et al. were able to accurately classify objects belonging to 1000 categories by using CNN trained on the ImageNet dataset [7]. The animal breed identification problem was tackled by many researchers using various methods, e.g., using CNN for dog and cat breeds classification [8], using deformable parts model for detecting wild kangaroos [9], using reinforcement learning technique called FCAN to detect dog breeds [10], using supervised clustering framework built from multi discriminative parts and expectation–maximization approach [11], using semi-supervised learning and multi-part convolutional neural network (MP-CNN) [12]. Yosinski et al. introduced transfer learning as a new technique for training deep CNNs [13]. CNN with transfer learning were used for dog breeds classification [14] and sheep breeds classification [15]. Animal breed

Fig. 1 Sample images from goat breed database

datasets have also been constructed for various uses including animal breed identification tasks [16, 17].

Image Capturing & Goat Image Database

Out of the many registered goat breeds present in India, the most commonly available six goat breeds have been selected for this experiment, such as JAKHRANA, BLACKBENGAL, SIROHI, BARBARI, BEETAL and JAMUNAPRI as shown in Fig. 1.

These pure breed goat images have been captured by Samsung A6 mobile device from the organized farms, namely; of 1) ICAR-Indian Veterinary Research Institute, Kolkata, West Bengal; 2) Uttarbanga Krishi Viswavidyalaya, Cooch Behar, West Bengal and 3) ICAR-Central Institute for Research on Goat, Uttar Pradesh.

The images have been captured by isolating a single goat and capturing its full contour and in some cases the front facing view has also been captured as shown in Fig. 2.

The animal has been placed in two different backgrounds for generating diversity in our images, such as restricted and unrestricted (Fig. 2). In restricted images the background of the image has been controlled by putting a green screen behind it, but in the unrestricted images the animal has been photographed naturally without restricting the background objects, such as in the grazing field, living sheds. For capturing photographs in both restricted and unrestricted environments, goats were not subjected to any distress condition which could violate welfare of farm animals.

Augmentation

By using different photographing techniques, at least 150 distinct images could be collected from each of the 6 breeds. But the number of original images varies widely, since every breed did not contain the same number of goats, few breeds like Sirohi and Jamunapari had only 50 to 60 goats while breeds like Beetal had more than 200 goats. Accuracy of a deep learning model can be largely attributed to the amount of training data and diversity of the training data. So, a good number of training data with large diversity is necessary to build a model with high accuracy. This disparity of inter breed training images has been solved by data augmentation techniques. Data augmentation is the process by which the diversity and amount of the training images can be increased significantly by applying some geometric as well as image processing techniques, without compromising the quality of the training database. Some of the techniques used for data augmentation used in

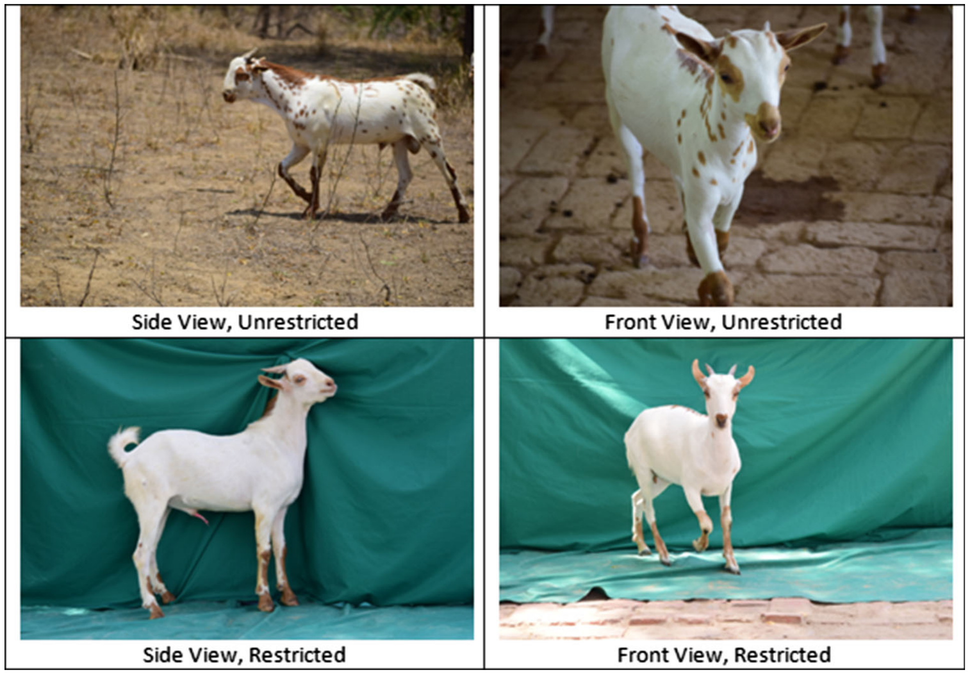


Fig. 2 Types of images in goat breed database

Fig. 3 Data augmentation

Table 1 Details of goat breed database

Breed name	No. of train images	Augmented	No. of test images	Total images
Barbari	280	No	30	310
Beetal	280	Yes	30	310
Blackbengal	280	Yes	30	310
Jakhrana	280	No	30	310
Jamunapari	280	Yes	30	310
Sirohi	200	Yes	25	225

this paper are rotation by arbitrary degrees, width and height translations and shearing transformations as shown in Fig. 3.

In real life a goat can deform in various ways, such as sitting, standing, curling up, etc., but humans can still identify it as a goat nonetheless. This is the main idea behind data augmentation that the same image can be purposefully deformed and fed to the training pipeline so that the model becomes resistant to the deformable nature of the object it's identifying.

Finally, all these images have been resized to (512 $\times$ 512) pixels and stored in the disk. The details of goat breed database are given in Table 1.

Convolutional Neural Network

A convolutional neural network (CNN) is an image classification algorithm. It takes a set of images as input (training images), uses those images to adjust its learnable weights and biases by calculating loss from the image ground truth labels and backpropagating through the whole network. In this way, a CNN is able to predict the type of the object present in a different set images by learning the important features available from the images in the training set, given both the set of images are similar in some way [18]. Any CNN model consists of the following layers: convolution layers, pooling layers and fully connected layers as shown in Fig. 4.

The number of each kind of layer, known as the depth of the CNN network may vary depending on the complexity of the problem and computational resources available. AlexNet, VGG16, GoogLeNet, ResNet are few CNN architectures used widely for image classification.

Convolutional Layer

The convolutional layer is the core building block of a Convolutional Neural Network that does most of the computational heavy lifting. The Convolutional layer's parameters consist of a set of learnable filters. Every filter is small spatially (along width and height), but extends through the full depth of the input volume. For example, a typical filter on a first layer of a CNN might have size $5 * 5 * 3$ (i.e., 5 pixels width and height, and have a depth of 3, the color channels). If the input image matrix is of dimension $a * b * c$ and filter matrix is of dimension $e * f * c$ then the dimension of output matrix will be $(a - e + 1) * (b - f + 1) * 1$ (Fig. 5).

During training, the program will slide (more precisely, convolve) each filter across the width and height of the input volume and compute dot products between the entries of the filter and the input at any position. As it slides the filter over the width and height of the input volume it will produce a 2-dimensional activation map that gives the responses of that filter at every spatial position. Intuitively, the network will learn filters that activate when they see some type of visual feature such as an edge of some orientation or a blob of some color on the first layer, or eventually entire honeycomb or wheel-like patterns on higher layers of the network. Now, a CNN will have an entire set of filters in each convolutional layer (e.g., 12

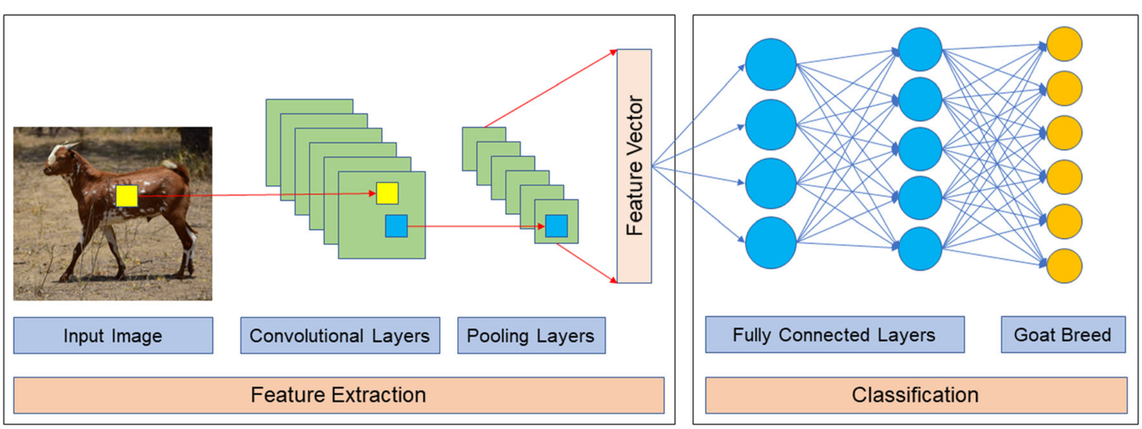


Fig. 4 CNN workflow

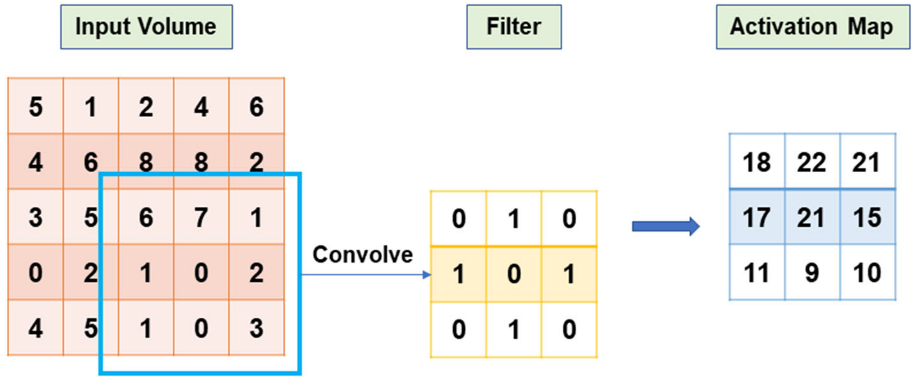


Fig. 5 Convolution operation

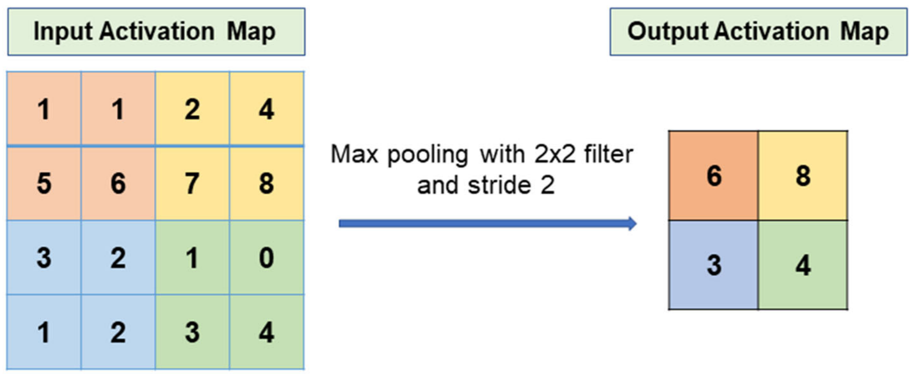


Fig. 6 Max pooling operation

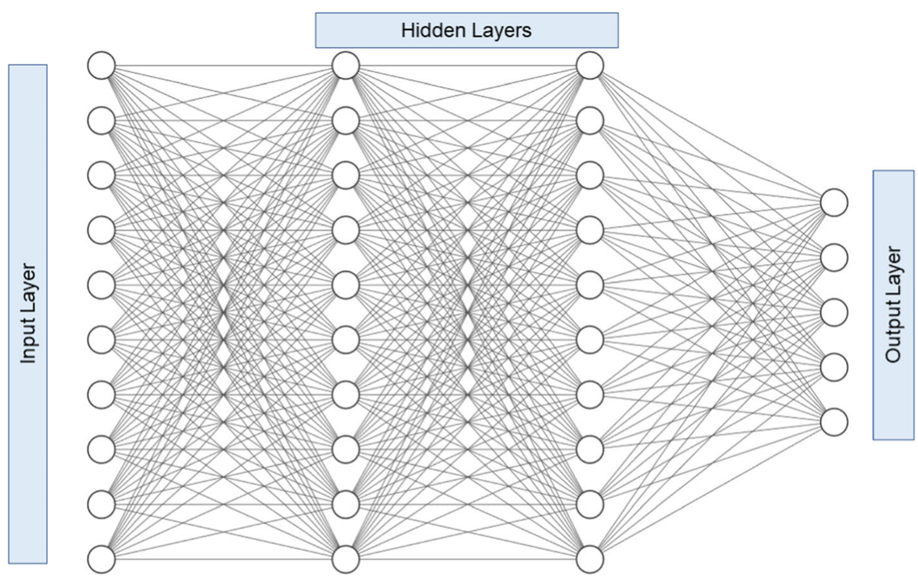


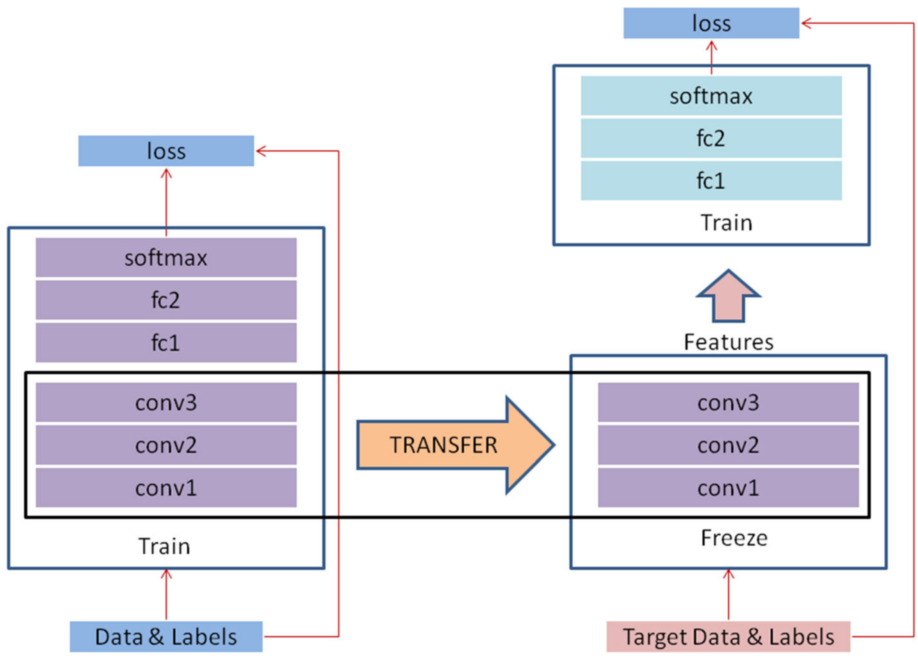
Fig. 7 Fully connected layer

filters), and each of them will produce a separate 2-dimensional activation map. These activation maps will then be stacked along the depth dimension and produce the output volume.

Pooling Layer

It is common to periodically insert a pooling layer in-between successive Convolutional layers in a CNN architecture. Its function is to progressively reduce the spatial

Fig. 8 Transfer learning workflow



size of the representation to reduce the number of parameters and computation in the network, and hence to also control overfitting. The pooling layer operates independently on every depth slice of the input and resizes it spatially, selecting the maximum value from within the filter. The most common form is a pooling layer with filters of size $2 * 2$ applied with a stride of 2 down samples every depth slice in the input by 2 along both width and height, discarding 75% of the activations. Every max pooling operation would in this case be taking a max over 4 numbers (Fig. 6). The depth dimension remains unchanged.

Fully Connected Layer

After creating feature vector by flattening the last convolutional feature map it is then passed through a dense neural network for classification (whether the input image is goat or cat or dog or tree or anything). This dense network is called fully connected layer. Neurons in a fully connected layer have full connections to all activations in the previous layer. Their activations can hence be computed with a matrix multiplication followed by a bias offset. The neuron computes the weighted sum of all inputs along with their connecting weights and biases. The result is fed into an activation function which is again fed into the next neurons of the next layer. Finally, the output is produced from output layers. The basic structure of a fully connected layer is depicted in Fig. 7 [19].

Transfer Learning

Transfer learning is the process by which the already learned weights and biases acquired by a model trained on a large standardized dataset called pre-trained model is reused [13]. This pre-trained model's few top-level layers are then retrained by freezing the weight and bias updates of the lower layers based on the problem space (Fig. 8).

The number of top layers retrained depends on the availability of training images and computation power available. For a very small number of training images only the fully connected layers are retrained. In this way the fully connected layers get trained according to the problem space where the old model acts as the feature extractor and reduces training time greatly but also provides high classification accuracy. This is also the go-to method for using large networks such as Inception, ResNet, Inception-ResNet, etc., which have up to 150 layers, because training such networks end-to-end is a very time-consuming task.

Performance Evaluation

Performance evaluation is an important aspect of the deep learning process. However, it is a complex task. It, therefore, needs to be conducted carefully in order for the application of deep learning to the respective field to be reliable. There are various evaluation metrics available for classification and localization [20–22]. In this paper, the primary objective is to train a deep-learning-based model which should be able to predict the correct breed of an animal from among multiple breeds. This falls under the multi-class classification problem. Such models can be evaluated while training using the validation set images and then when the trained model is used for actual prediction using the test set images.

Model evaluation while training is done based on the loss and accuracy values obtained after each iteration of training on the train set images compared with the validation set images. If training is done properly, the loss should gradually decrease and become close to zero and at the same time the accuracy should increase and become close to one. A validation set, as the name suggests is used to check whether the model is actually able to learn from the train set and apply its knowledge as expected. If the trained model performs better on the validation set, it is said to be perfectly trained. Plots of train vs. validation loss and train vs. validation accuracy are frequently used to perform such evaluations.

For evaluating a model's predictive performance different set of metrics are used. For any model, its predictions can be categorized into three broad categories, namely true positive (TP) predictions, false positive (FP) predictions and false negative (FN) predictions. Determining the number of such predictions varies with the evaluation conditions used. In case of multi-class classification, true negative (TN) are not considered, since TN predictions for a particular class will belong to TP, FP or FN prediction for other classes.

For a multi-class classification environment, the model should be able to predict the correct breed of the animal with certain level of confidence. For such cases, a value called confidence score is used. Based on the confidence score threshold with which the prediction is done, TP, FN and FP can be determined as given in Table 2. The confidence score is actually the minimum accuracy with which the model is able to say that the input belongs to a certain class.

Based on TP, FN and FP values, three very important performance evaluation metrics can be calculated which have been used in this thesis to compare the model performance.

Table 2 Determining TP, FN & FP for classification with confidence score

Category	Conditions
True positive (TP)	Confidence Score $\geq$ Confidence Score$_{\text{threshold}}$ & Actual Class = Predicted Class
False positive (FP)	Confidence Score $\geq$ Confidence Score$_{\text{threshold}}$ & Actual Class $\neq$ Predicted Class
False negative (FN)	Confidence Score < Confidence Score$_{\text{threshold}}$

1. *Precision* Precision answers the question: Out of the items that the model predicted to be true, how many are actually true?

   It is calculated as: Precision $= \frac{TP}{TP+FP}$.

2. *Recall* Recall answers the question: Out of all the items that are true, how many are found to be true by the model?

   It is calculated as: Recall $= \frac{TP}{TP + FN}$.

3. *Accuracy* Accuracy measures the number of true predictions out of all the predictions done by the model.

   It is calculated as: Accuracy $= \frac{TP}{Total\ Test\ Cases}$.

A model is said to be perfect if it is able to predict all the test set images correctly. For example, if the test set contains x number of images, for a perfect model, $TP = x$, $FP = 0$ & $FN = 0$. Therefore, Precision $= 100\%$, Recall $= 100\%$ & Accuracy $= 100\%$. So, to build a good model, the objective is to reduce the number of FP and FN predictions and increase TP predictions so that all the three metrics (precision, recall and accuracy) acquires values close to 100%.

Proposed Method

In this paper, an algorithm is proposed for classification of six goat breeds from their images. The algorithm consists of six modules; starting from capturing images to their classification and they work as layered structure as shown in Fig. 9.

The Image Capturing module and Image Database creation module has already been discussed. The CNN Model selection module is needed because, even after in-depth literature review, it could not be established which deep convolutional model will perform best for any set of images. That means for different data, there may be different deep convolutional models that will fit the data in an optimized manner. For this disparity, two different deep convolutional models namely VGG-16 [23] and Inception-v3 [24] were trained on the goat breed database using transfer learning and performance compared. For the sake of comparison, all the model hyperparameters for both the models were kept identical. The models were pre-trained on ImageNet data [25] trained for 100 epochs with the RMSprop [26] optimization algorithm and a fixed learning rate of 0.0001.

To avoid overfitting and under-fitting problem, the training images are divided into training and validation image sets. The training set is used to train the model, while the validation set is only used to evaluate the model's

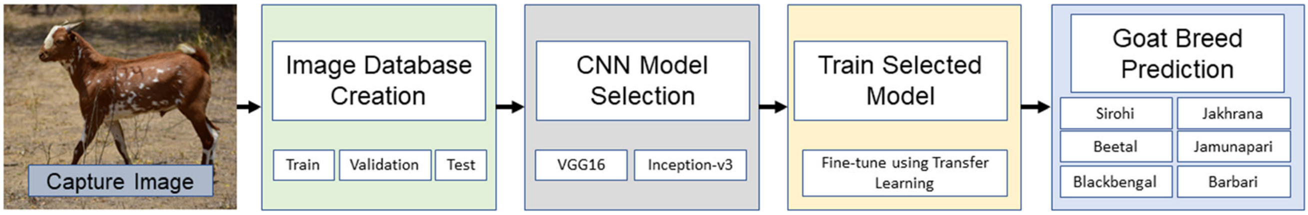


Fig. 9 Proposed algorithm for goat breed prediction

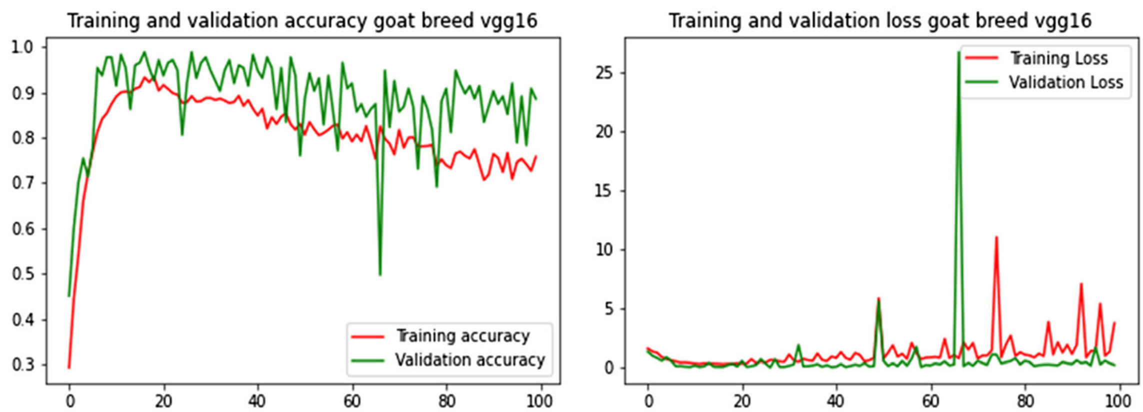


Fig. 10 Accuracy & loss graph of training & validation for VGG-16

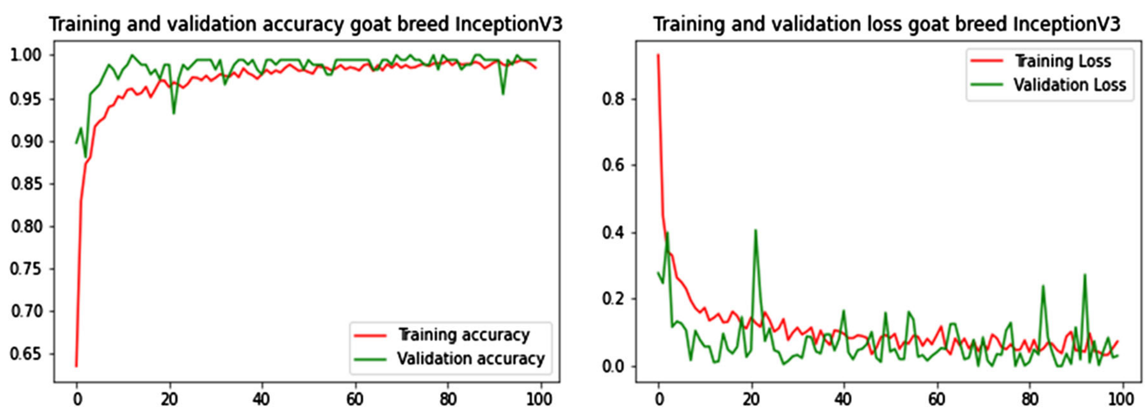


Fig. 11 Accuracy & loss graph of training & validation for Inception-v3

Table 3 Comparison of training statistics

Model	No. of layers	Hyperparameters	Training accuracy	Validation accuracy	Training loss	Validation loss
VGG-16	16	Learning rate: 0.0001 Iterations: 100	0.7575	0.8857	3.7880	0.2374
Inception-v3	42		0.9850	0.9943	0.0724	0.0307

performance. The goat breed database was similarly partitioned into two sets namely the training set containing 90% of the total images and the validation set with 10% of the total images. The good fit is said to have been obtained when the accuracy of validation image set is very close to training image set. This indicates that the model is trained

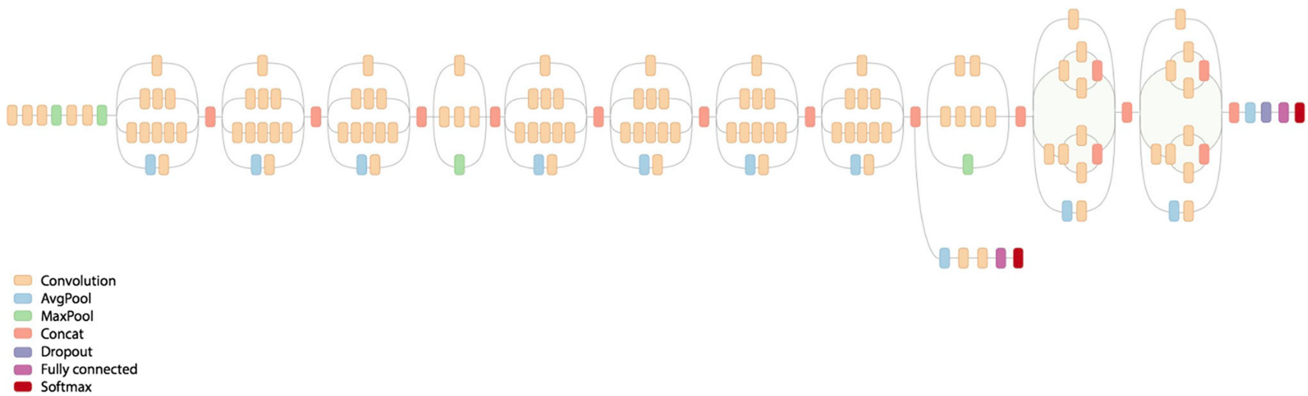


Fig. 12 Inception-v3 architecture

Table 4 Goat breed prediction result at various confidence thresholds

Confidence score	Breed	Total cases	TP	FN	FP	Precision	Recall
0.92	Barbari	30	29	0	1	0.96	1.00
	Beetal	30	30	0	0	1.00	1.00
	Blackbengal	30	29	1	0	1.00	0.96
	Jakhrana	30	29	0	1	0.96	1.00
	Jamunapari	30	30	0	0	1.00	1.00
	Sirohi	25	24	0	1	0.96	1.00
0.93	Barbari	30	28	1	1	0.96	0.96
	Beetal	30	30	0	0	1.00	1.00
	Blackbengal	30	29	1	0	1.00	0.96
	Jakhrana	30	29	1	0	1.00	0.96
	Jamunapari	30	30	0	0	1.00	1.00
	Sirohi	25	24	1	0	1.00	0.96
0.95	Barbari	30	28	2	0	1.00	0.93
	Beetal	30	30	0	0	1.00	1.00
	Blackbengal	30	28	2	0	1.00	0.93
	Jakhrana	30	29	1	0	1.00	0.96
	Jamunapari	30	30	0	0	1.00	1.00
	Sirohi	25	24	1	0	1.00	0.96
0.96	Barbari	30	28	2	0	1.00	0.93
	Beetal	30	29	1	0	1.00	0.96
	Blackbengal	30	28	2	0	1.00	0.93
	Jakhrana	30	29	1	0	1.00	0.96
	Jamunapari	30	30	0	0	1.00	1.00
	Sirohi	25	22	3	0	1.00	0.88
0.97	Barbari	30	28	2	0	1.00	0.93
	Beetal	30	29	1	0	1.00	0.96
	Blackbengal	30	28	2	0	1.00	0.93
	Jakhrana	30	29	1	0	1.00	0.96
	Jamunapari	30	28	2	0	1.00	0.93
	Sirohi	25	22	3	0	1.00	0.88

properly. The training versus validation accuracy and loss graphs for VGG-16 is given in Fig. 10 and for Inception-v3 is given in Fig. 11.

The models were compared based on the convergence trend of training and validation in similar number of iterations. In Fig. 11, the validation accuracy and loss converge with training accuracy and loss, which indicates that the Inception-v3 produced good fit on Goat Image dataset. Similarly, looking at Fig. 10, the training and validation accuracy and loss both do not converge, also the overall accuracy for VGG-16 is lower than Inception-v3 (Table 3).

As a result, Inception-v3 model was selected as the preferred model to predict the goat breeds from goat images. The overall architecture of Inception-v3 network is given in Fig. 12 [27].

It is 42 layers deep and has only 7 million trainable parameters compared to VGG16 which is only 16 layers deep and has 180 million trainable parameters. This drastic reduction in number of parameters along with using suitably factorized convolutions and aggressive regularization techniques enables Inception-v3 to be computationally efficient to train and produce much higher classification accuracies.

In the train selected model module, Inception-v3 model has been built and implemented using Keras API [28] and trained on a machine with 3.5 GHz Xeon Processor with 64 GB of Ram and dual NVidia Quadro P4000 GPU's.

Result & Discussion

In the Goat Breed Prediction module, for goat breed prediction, 175 images spread over six breeds have been used and have been tested at different confidence thresholds as given in Table 4.

From these prediction results precision, recall and accuracy metrics have been computed (Table 5). The

Table 5 Goat Breed Prediction Metrics

Confidence score	Average precision %	Average recall %	Accuracy %
0.92	98.00	99.33	97.71
0.93	99.33	97.33	97.14
0.95	100.00	96.33	96.57
0.96	100.00	94.33	94.85
0.97	100.00	93.16	93.71

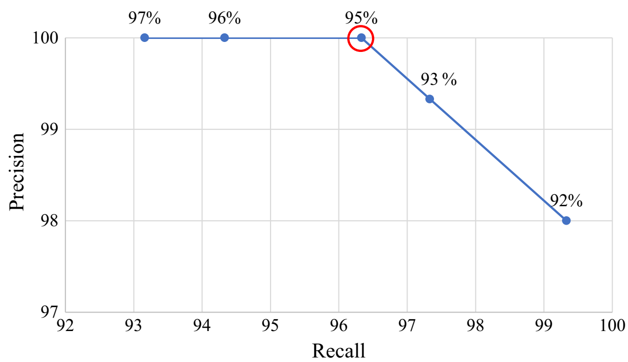


Fig. 13 Precision vs. recall for goat breed prediction

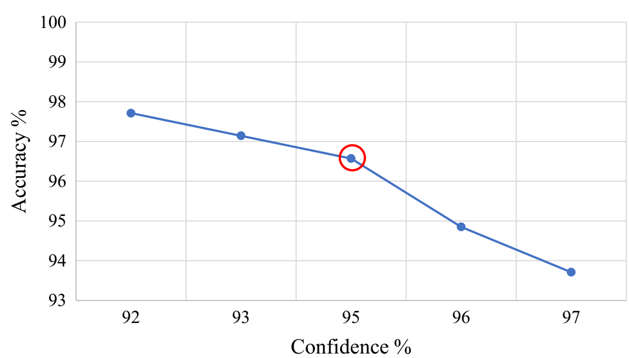


Fig. 14 Goat breed prediction accuracy at different confidence thresholds

objective is to find the maximum confidence level for which all the metrics have maximum values.

Plotting the precision vs. accuracy curve at different confidence intervals (Fig. 13), we can observe that with decreasing confidence threshold, precision decreases but recall increases. Therefore, a trade-odd between confidence, precision and recall is needed. From the plot, the confidence level of 95% is found to be maximum confidence, for which both precision (100%) and recall (96.33%) are high.

Table 6 Performance of species categorization

Method	Species	Accuracy (%)
Selective pooling of local visual descriptors with SVM classifier [29]	Dog	52.00
Classification by extracting SIFT descriptors [8]	Cat & dog	59.00
Multiple kernel framework with SVM classifier [30]	Flower	72.80
Part-based R-CNNs [31]	Bird	76.37
Object segmentation through propagation approach [32]	Bird	80.66
Deformable part model (DPM) [9]	Kangaroo	84.25
Reinforcement learning technique called FCAN [10]	Dog	88.90
Fine-tuning of pre-trained deep learning models [14]	Dog	89.66
A mix of CNN, SIFT and SVM [33]	Dog	90.00
Fine-tuning of pre-trained deep learning models [34]	Dog	90.69
Fine-tuning VGG-16 model [15]	Sheep	95.80
Fine-tuning inception-v3 (This paper)	Goat	96.57

Plotting the prediction accuracy against the confidence levels (Fig. 14), it can be observed that with increasing confidence, accuracy decreases.

As a result, the confidence level of 95% is chosen and the prediction accuracy (96.57%) for that confidence is taken as the goat breed prediction accuracy for this model. This prediction accuracy has been compared with other research done for animal breed identification as shown in Table 6. From the comparison it is established that goat breed prediction by fine-tuning Inception-v3 produces better results compared to others.

Conclusion & Future Work

In this paper, the authors proposed an algorithm for classification of six goat breeds from their images. The algorithm simulates human visual recognition system using convolutional neural network and it successfully classified six different goat breeds from their images with 96.57% accuracy at 95% confidence level. In this effort, image database for the six Indian goat breeds, named Blackbengal, Beetal, Jamunapari, Barbari, Jakhrana and Sirohi was established. The authors are hopeful that the trained model can be used as ready to use technology to help recognizing breed identity of a goat from its image. More breeds may be tested and more new models may be developed in future.

Acknowledgements The authors are thankful to the AG & Food division of Information Technology Research Academy (ITRA), Media Lab Asia, New Delhi, India, for grant in aid for conducting the research. The authors gratefully acknowledged unconditional support from the Director, ICAR- CIRG Makhdoom, UP and the Joint Director of ICAR-IVRI Eastern Regional Station, Kolkata. For constant encouragement, constructive criticism and scientific input we all acknowledge with reverence the support from Dr. Amitabha Bandyopadhyay, Senior Consultant of ITRA Ag & Food. The Investigating team of the partner institutes–ICAR Research Complex for NEH region, Barapani and Indian Institute of Technology, Guwahati are duly acknowledged for their support. The authors are also thankful to Dr. Sourabh Kumar Das, Principal, Kalyani Government Engineering College for his continuous support.

References

1. H. Wang, B. Raj, On the origin of deep learning. *arXiv preprint* arXiv:1702.07800, 2017
2. A. Hailu, Breed characterization: tools and their applications. Open Access Libr J **2**(4), 1 (2015)
3. W. Peng, H. Yang, K. Cai, L. Zhou, Z. Tan and K. Wu, Molecular identification of the Danzhou chicken breed in China using DNA barcoding. Mitochondrial DNA Part B **4**(2), 2459–2463 (2019)
4. M. Teweldemedhn, M. Selam, Characterization of Begait cattle using morphometric and qualitative traits in Western Zone of Tigray, Ethiopia. Int J Livestock Prod **11**(1):21–33 (2020)
5. J. Sunder, A. Kundu, M. Kundu, T. Sujatha, A. K. De, Farming practices and morphometric characterization of Andaman Local Goat. Ind J Anim Res 53(8), 1097–1103 (2019)
6. S. Kumar, S.K. Singh, Visual animal biometrics: survey. IET Biometrics **6**(3), 139–156 (2016)
7. A. Krizhevsky, I. Sutskever, G. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012
8. O. M. Parkhi, A. Vedaldi, A. Zisserman, C. V. Jawahar, Cats and dogs. in *IEEE conference on computer vision and pattern recognition*, 2012
9. T. Zhang, A. Wiliem, G. Hemsony, B. C. Lovell, Detecting kangaroos in the wild: the first step towards automated animal surveillance. in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015
10. X. Liu, T. Xia, J. Wang, Y. Lin, Fully convolutional attention localization networks: efficient attention localization for fine-grained recognition," *arXiv preprint* arXiv:1603.06765, vol. 1, no. 2, pp. 4–4, 2016
11. D. Sundaram, A. Loganathan, A new supervised clustering framework using multi discriminative parts and expectation–maximization approach for a fine-grained animal breed classification (SC-MPEM), Neural Processing Letters, pp. 1–40, 2020
12. S. D. l. A. L. Meena, An efficient framework for animal breeds classification using semi-supervised learning and multi-part convolutional neural network (MP-CNN). IEEE Access **7**, 151783–151802 (2019)
13. J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks?, in *Advances in neural information processing systems*, 2014
14. A. Ayanzadeh, S. Vahidnia, A modified deep neural networks for dog breeds identification," *Preprints.org,* 2018
15. S.A. Jwade, A. Guzzomi, A. Mian, On farm automatic sheep breed classification using deep learning. Comput Electr Agric **167**, 105055 (2019)
16. U. A. Khan, S. M. U. Din, S. A. Lashari, M. A. Saare, M. Ilyas, Cowbree: A novel dataset for fine-grained visual categorization. Bull Electric Eng Inf **9**(5), 1882–1889 (2020)
17. A. Khosla, N. Jayadevaprakash, B. Yao, F.-F. Li, Novel dataset for fine-grained image categorization: Stanford dogs, vol. 2, no. 1, 2011
18. L. Fei-Fei, K. Andrej, J. Justin, CS231n: convolutional neural networks for visual recognition 2015. [Online]. Available: http://cs231n.stanford.edu
19. I. Goodfellow, B. Yoshua, C. Aaron, *Deep Learning* (MIT Press, Cambridge, 2016)
20. E. Mark, V.G. Luc, K.I.W. Christopher, W. John, Z. Andrew, The pascal visual object classes (voc) challenge. Int J Comput Vis **88**(2), 303–338 (2010)
21. J. Nathalie, S. Mohak, Performance evaluation in machine learning. in *Machine Learning in Radiation Oncology,* pp. 41–56, 2015
22. L. Tsung-Yi, M. Michael, B. Serge, B. Lubomir, G. Ross, H. James, P. Pietro, R. Deva, D. Piotr, Z. C. Lawrence, Microsoft coco: Common objects in context, in *European conference on computer vision*, 2014
23. K. Simonyan, Z. Andrew, Very deep convolutional networks for large-scale image recognition, *arXiv preprint* arXiv:1409.1556, 2014
24. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016
25. D. Jia, D. Wei, S. Richard, L. Li-Jia, L. Kai, F.-F. Li, ImageNet: a large-scale hierarchical image database. in *2009 IEEE conference on computer vision and pattern recognition*, 2009
26. T. Tijmen, H. Geoffrey, Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. COURSERA Neural Netw Mach Learn **4**(2), 26–31 (2012)
27. S. Jon, Train your own image classifier with inception in tensorFlow, Google AI Blog, 2016. [Online]. Available: https://ai.googleblog.com/2016/03/train-your-own-image-classifier-with.html. Accessed 10 June 2020
28. C. Francois, Keras. 2015. [Online]. Available: https://keras.io
29. G. Chen, J. Yang, H. Jin, E. Shechtman, J. Brandt, T. X. Han, Selective pooling vector for fine-grained recognition. in *IEEE Winter Conference on Applications of Computer Vision*, 2015
30. M. E. Nilsback, A. Zisserman, Automated flower classification over a large number of classes. in *Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, 2008
31. N. Zhang, J. Donahue, R. Girshick, T. Darrell, Part-based R-CNNs for fine-grained category detection. in *European conference on computer vision*, 2014
32. A. Angelova, S. Zhu, Efficient object detection and segmentation for fine-grained recognition. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013
33. W. LaRow, B. Mittl, V. Singh, Dog breed identification. *Network,* 2016
34. Z. Ráduly, C. Sulyok, Z. Vadászi, A. Zölde, Dog breed identification using deep learning. in *IEEE 16*th *International Symposium on Intelligent Systems and Informatics (SISY)*, 2018

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.