Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 16s (2025)
https://www.jneonatalsurg.com

Cataloguing of Animal Species Utilizing Convolutional Neural Networks

Dr.S. Saranya[1], Dr.D. Priyadharsini[2], Dr.S. Sasikala[3]

[1]Assistant professor, Department of Computer Science with Cognitive System and AIML, Hindusthan College of arts and science

[2]Assistant professor, Department of Computer Science with Cognitive System and AIML Hindusthan College of arts and science

[3]Head & Professor, Department of Computer Science with Cognitive System and AIML Hindusthan College of arts and science

Cite this paper as: Dr.S. Saranya, Dr.D. Priyadharsini, Dr.S. Sasikala, (2025) Cataloguing of Animal Species Utilizing Convolutional Neural Networks. *Journal of Neonatal Surgery*, 14 (16s), 425-436.

ABSTRACT

The growth of urban areas in contemporary times has led to significant habitat displacement in forested regions[1]. Consequently, wild animals are compelled to enter human settlements, which often disrupt their natural behaviors. Frequently, the search for food drives these animals to venture into populated areas. This situation poses a real threat to humans who may inadvertently encounter these animals when they are in a heightened state of aggression. Therefore, there is a pressing need for the identification of wild animals at the periphery of human communities adjacent to natural habitats[2]. An effective, reliable, and robust early warning system would greatly reduce the risk of deadly human- animal conflicts, thereby safeguarding both human lives and protecting endangered species.Moreover, such a system would also be useful in wildlife sanctuaries and biosphere reserves to monitor the movement of animals at the border areas of such establishments which have often proved difficult to control[4]. The usage of technology and robust cameras is not an alien concept in most major biosphere reserves and national parks around the world. Although there has been a considerable amount of progress, software-based tools have not been explored to a satisfactory extent in these use cases.[3]

Computer vision has the ability to transform the tracking and monitoring process with the accuracy that its components and supporting techniques provide[5][8]. The automation-augmented reduction of man-hours invested in searching for and tracking wild animals is perhaps the biggest potential boon that computer vision can provide. The pre-processing involved in the application of computer vision algorithms is often under-documented although it plays a key role in the success of the algorithm[6]. A deep understanding of the nature of the inputs is necessary to make appropriate changes at crucial junctures of processing to meet the often- convoluted criteria required by complicated deep learning algorithms. Transforming the images is invariably necessitated due to the erratic nature of real-world data feeds[4]. The absence of an artificial synthesis element in the generation of inputs via raw camera stills adds to the intricacies involved in the image processing component. Many researchers have studied on the image classification. The purpose of the study is to establish a model that can realize the animal image classifing by using CNN. In this paper, first of all, a suitable data set is collected for the research[5]. Secondly, the parameters and layers of the neural network are chosen. Finally, a suitable criterion is established for evaluating the model so as to find the most suitable model to solve the research problem[8].

1. INTRODUCTION

Animal identification using a means of marking is a process done to identify and track specific animals. It is done for a variety of reasons including verification of ownership, biosecurity control, and tracking for research or agricultural purposes. Animal identification is known as ‘the combination of the identification and registration of an animal individually, with a unique identifier, or collectively by its epidemiological unit or group, with a unique group identifier.

However, there are other benefits of animal identification, such as accurate breeding histories of individual animals, significant contributions to breeding and genetic diversity management programmes and in the implementation of targeted biosecurity measures (Richard, n.d).

A Convolutional Neural Network (CNN) is a type of deep learning algorithm specifically designed for image processing and recognition tasks. Compared to alternative classification models,

CNNs require less preprocessing as they can

automatically learn hierarchical feature representations from raw input images. They excel at assigning importance to various objects and features within the images through convolutional layers, which apply filters to detect local patterns[1]. The connectivity pattern in CNNs is inspired by the visual cortex in the human brain, where neurons respond to specific regions or receptive fields in the visual space. This architecture enables CNNs to effectively capture spatial relationships and patterns in images. By stacking multiple convolutional and pooling layers, CNNs can learn increasingly complex features, leading to high accuracy in tasks like image classification, object detection, and segmentation.

2. LITERATURE REVIEW

Most of the literature on FGC, have either used annotations for either objects or the parts for extracting the discriminative features. Hierarchical Part Matching (HPM) [4], Part-based One-vs.- One Features (POOF) [5], Pose Normalized Deep Convolutional Nets (PN-DCN) [6], and Part based R-CNN [7] utilizes annotations on both object and part level for both train and test data. However, annotations are expensive and human-intensive. In particular, part annotations are tedious and prone to error. Deep Localization, Alignment and Classification (Deep-LAC) [8], and Part-Stacked CNN (PS-CNN) [9] proposed models for fine-grained visual categorization using annotations for object localization in both train and test images. Coarse-to-fine [10], Webly-supervised [11], and PG Alignment [1] used hand- crafted SIFT features and annotations on the object level.

Xie et al. introduced InterActive [3], an innovative algorithm designed to assess neuron attention and enhance the performance of low-level neurons for improved classification outcomes. In contrast, TL Atten [4] overlooked the spatial relationships between object parts and among the parts themselves. Simon and Rodner [5] developed a Constellation of Neural Activations (CAN) method to identify discriminative parts without relying on bounding boxes; however, their approach did not initially localize the object. Lin et al. [16] presented the Bilinear CNN, which integrates two convolutional neural networks for feature extraction. Nonetheless, this model does not incorporate any annotations, unlike Fused One-vs-All Features (FOAF) [7] and Dense Graph Matching (DGM) [18]..

Recently, Zheng *et al.* [9] proposed a novel Multi-Attention Convolutional Neural Network (MA- CNN) for fine-grained classification problems. The network consists of convolutional layers, channel grouping and a sub-network for part classification. One other common problem found in existing

FGC works is that they do not consider the relationship among the parts and the relationship between the features extracted by different parts [2]. In yet another case, when the object is localized before choosing the parts, then the spatial relationship between the object and the parts should not be ignored [21. The spatial relationship among the parts and between the object and the parts are highly useful in extracting discriminating parts, which makes the classification easier.

Generally, we face two common problems when objects are localized before discriminative components are chosen. The first problem is that the background of the localized objects is sometimes bigger than the object itself. The second problem is that there is a lot of overlap between the object and the backdrop, which makes information redundant. In order to overcome these obstacles, we present a brand-new Multi Part Convolutional Neural Network (MP CNN) that combines part selection and object localization models without the need for annotations. Animal picture classification has been the subject of numerous studies, all of which have produced datasets with differing degrees of accuracy.

Using deep neural networks on the Snapshot Serengeti dataset, Norouzzadeh [24] was able to categorize 48 mammal species from Tanzania's Serengeti National Park with an accuracy of almost 92%. By classifying 20 distinct animal species using a very deep convolutional neural network, Villa et al. [25] considerably improved the accuracy for the same dataset. They obtained an accuracy of roughly 88.9% in the Top-1 category and 98.1% in the Top-5 category by putting a residual network (ResNet) architecture into practice. Although a number of research [24], [26] have demonstrated impressive accuracy in recognizing animals from camera-trap photographs, their algorithms were only able to recognize a small number of images and relied on manually created feature extraction. Despite being an essential stage in image analysis, feature extraction has some drawbacks.

3. METHODOLOGY OF IDENTIFICATION:

Below are several commonly known methods of animal identification.

Electronic Animal Identification Devices: These are electronic tags containing radio frequency identifiers (RFID) that emit a signal when stimulated by an appropriate electronic reader. The RFID consists of a microchip and a coiled copper antenna (Shanahan et al., 2009). Although it is not necessary for farmers to possess readers, there can be labor cost savings in their use due to reduced time for reading and recording identification numbers; reader use may also improve accuracy[5].

Tattoo for animal identification: Tattoo is an animal identification method that is mostly used on pigs, particularly white-skinned breeds, although green ink can be used for dark-skinned breeds. Tattoos have also been proposed as a means to identify poultry; the tattoo would be applied under the wing where there are fewer feathers[6].

Ear Notching & Ear Clipping for animal identification: Ear notching is not an approved means of identification but may be used for on-farm management of pigs as an alternative to tattooing or ear tags: a special tool snips out a small piece of the ear flap to leave a permanent V-shaped notch[8]. Notches could be used (but not as an official mark) to identify a herd number on one ear and an individual on the other, or a litter number on one ear and an individual on the other[7].

There is a risk of infection in newly notched ears and the welfare concerns are greater than for

some other identification methods. Ear Clipping has a similar procedure to ear notching but a more frequently used term when used for livestock other than pigs.

Ear Tagging for animal identification: Ear tags are by far the most common method of marking cattle, sheep and goats and are sometimes used for pigs. Ear tags may be electronic or non- electronic and there are numerous designs, but all entail the piercing of the ear flap. Most tags are now plastic (especially for electronic tags for which metal would interfere with the signal from the RFID)[5], but some non-electronic ear tags are metal. Apart from the initial puncturing of the ear flap (with possible introduction of infectious agents), the welfare concerns of ear tags are that some 'loop' types, if incorrectly inserted, may limit the growth of the ear and that all types may catch on fences etc. and be torn out of the ear[4].

Dewlap Tags for animal identification: Dewlap tags are inserted into the dewlap over the brisket, where it is claimed there are few nerves in the skin. A supplied hole-punch is used to make a hole in the skin before passing the hasp of the tag through the dewlap[4]; the tag is then passed over the hasp before the ends of the hasp are bent to retain the tag. Advantages claimed (by the suppliers) for dewlap tags over ear tags are that they are always visible (from in front of the animal) and the right way up, and that they do not snag on fences etc.

Pastern (Leg) Bands for animal identification: Pastern bands are a non- invasive method in which a band is fitted around the animal's lower leg; the pastern band may be electronic or non- electronic. The bands are not used anymore because they cause inflammation.

Ruminal Bolus for animal identification: A ruminal bolus is a bullet-shaped, high density container, usually ceramic, containing an electronic identifier; the bolus is designed to lodge in the rumen of the animal. Ruminal boluses were used as a means of identifying sheep during the National Scrapie Programme (NSP), but are not widely used for general identification as an ear tag is also needed as a visual identifier[8]. If this secondary visual identifier is lost, an electronic reader must be used to identify the animal. In addition, boluses are sometimes regurgitated (although ear tags and pastern bands can also be lost).

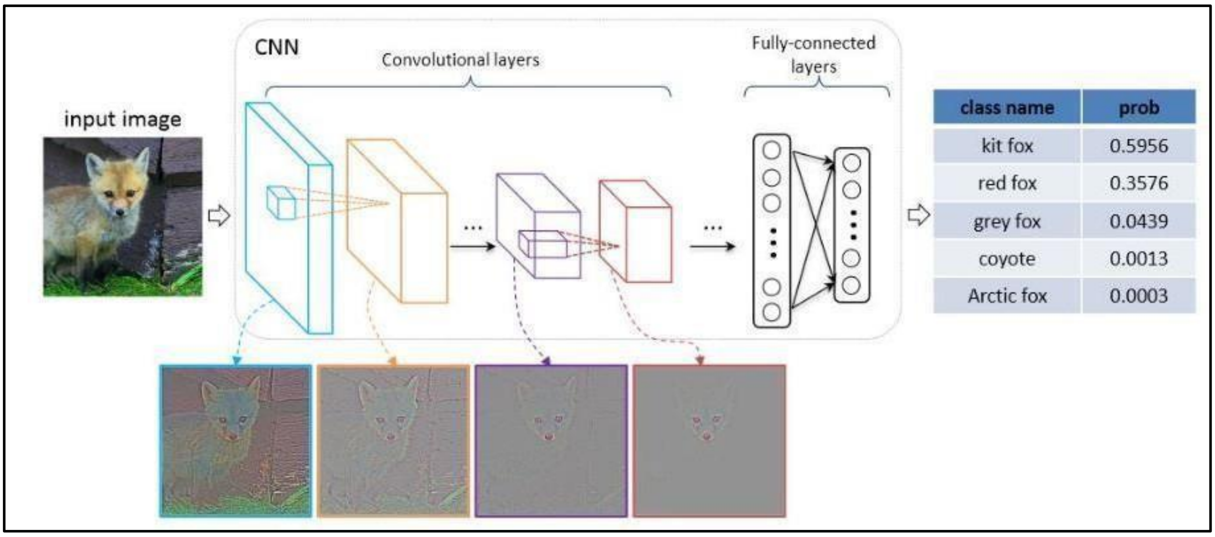
Freeze Branding for animal identification: Freeze branding is permitted for cattle as well as equines and is now more common than hot branding. The branding 'iron' (copper and brass are preferred metals) is cooled in liquid nitrogen, dry ice or similar coolant. When sufficiently cold the iron is applied to a shaved area of the animal's skin and held there for 15-45 seconds, the longer time on light coloured animals[4].

Micro-chipping for animal identification: Micro-chipping is the insertion of an electronic transponder directly under the skin of the animal, where it remains a permanent identifier that can be read with an appropriate electronic reader. Micro-chipping is widely used for companion animals such as dogs, but it is also common in equines.There is a risk that the microchip will migrate, making it more difficult to locate. For this reason microchipping is not commonly used for animals that may be slaughtered for human consumption: if the microchip cannot be quickly recovered it will interfere

with abattoir throughput or, if not recovered at all, it may pose a risk if ingested by the consumer[5].

VGG is a classical convolutional neural network architecture. It was based on an analysis of how to increase the depth of such networks. The network utilises small 3 x 3 filters. Otherwise the network is characterised by its simplicity: the only other components being pooling layers and a fully connected layer[10][3]. VGG neural network accepts a 224×224 pixel RGB image as input. To keep the input image size consistent for the ImageNet competition, the authors clipped out the middle 224×224 patch in each image. VGG's hidden layers all use ReLU, an AlexNet invention that significantly reduces training time. Local Response Normalisation (LRN) is not used by VGG since it increases memory usage and training time without improving accuracy. VGG contains three completely connected layers, the first two of which each has 4096 channels and the third of which has 1000 channels, one for each class.

VGG Architecture VGGNets are based on the most essential features of convolutional neural networks (CNN). The following graphic shows the basic concept of how a CNN works:

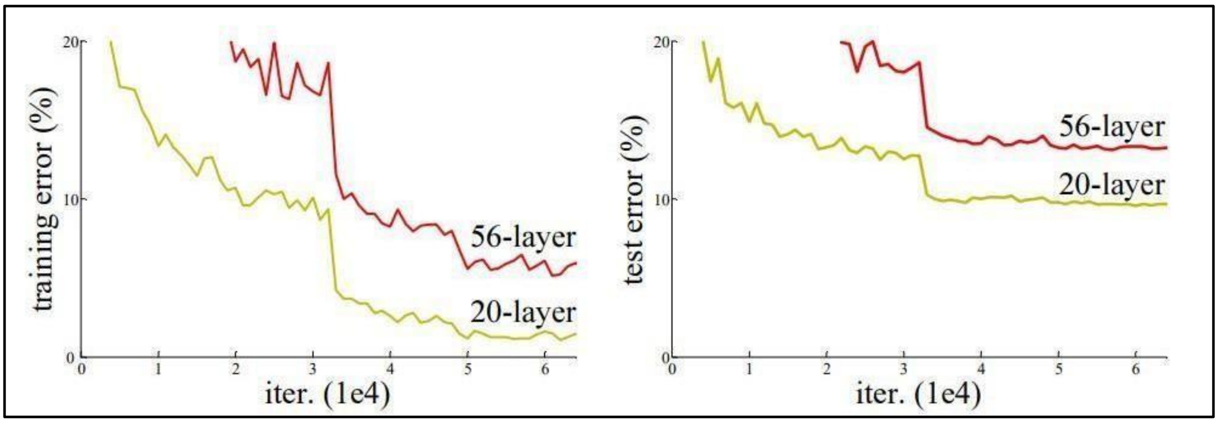


VGG stands for Visual Geometry Group; it is a standard deep Convolutional Neural Network (CNN) architecture with multiple layers. The “deep” refers to the number of layers with VGG-16 or VGG- 19 consisting of 16 and 19 convolutional layers[4]. The VGG architecture is the basis of ground- breaking object recognition models. Developed as a deep neural network, the VGGNet also surpasses baselines on many tasks and datasets beyond ImageNet. Moreover, it is now still one of the most popular image recognition architectures.

Complexity and challenges:

The number of filters that we can use doubles on every step or through every stack of the convolution layer. This is a major principle used to design the architecture of the VGG16 network. One of the crucial downsides of the VGG16 network is that it is a huge network, which means that it takes more time to train its parameters[5]. Because of its depth and number of fully connected layers, the VGG16 model is more than 533MB. This makes implementing a VGG network a time- consuming task. The VGG16 model is used in several deep learning image classification problems, but smaller network architectures such as GoogLeNet and SqueezeNet are often preferable. In any case, the VGGNet is a great building block for learning purposes as it is straightforward to implement[6].

ResNet provides an innovative solution to the vanishing gradient problem, known as “skip connections”. ResNet stacks multiple identity mappings (convolutional layers that do nothing at first), skips those layers, and reuses the activations of the previous layer. Skipping speeds up initial training by compressing the network into fewer layers.



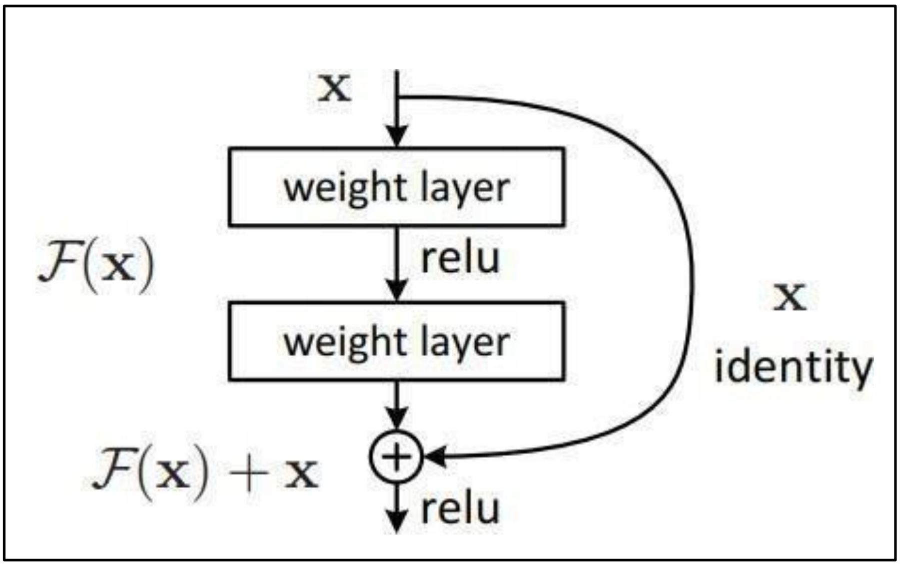
Comparison of 20-layer vs 56-layer architecture

In the above plot, we can observe that a 56-layer CNN gives more error rate on both training and testing dataset than a 20-layer CNN architecture. After analysing more on error rate the authors were able to reach the conclusion that it is caused by vanishing/exploding gradient. ResNet, which was proposed in 2015 by researchers at Microsoft Research, introduced a new architecture called Residual Network[3].

Residual Network: In order to solve the problem of the vanishing/exploding gradient, this architecture introduced the concept called Residual Blocks. In this network, we use a technique called skip connections. The skip connection connects activations of a layer to further layers by skipping some layers in between. This forms a residual block. Resnets are made by stacking these residual blocks together[3].

The approach behind this network is instead of layers learning the underlying mapping, we allow the network to fit the residual mapping. So, instead of say H(x), initial mapping, let the network fit,

$F(x) := H(x) - x$ which gives $H(x) := F(x) + x$.

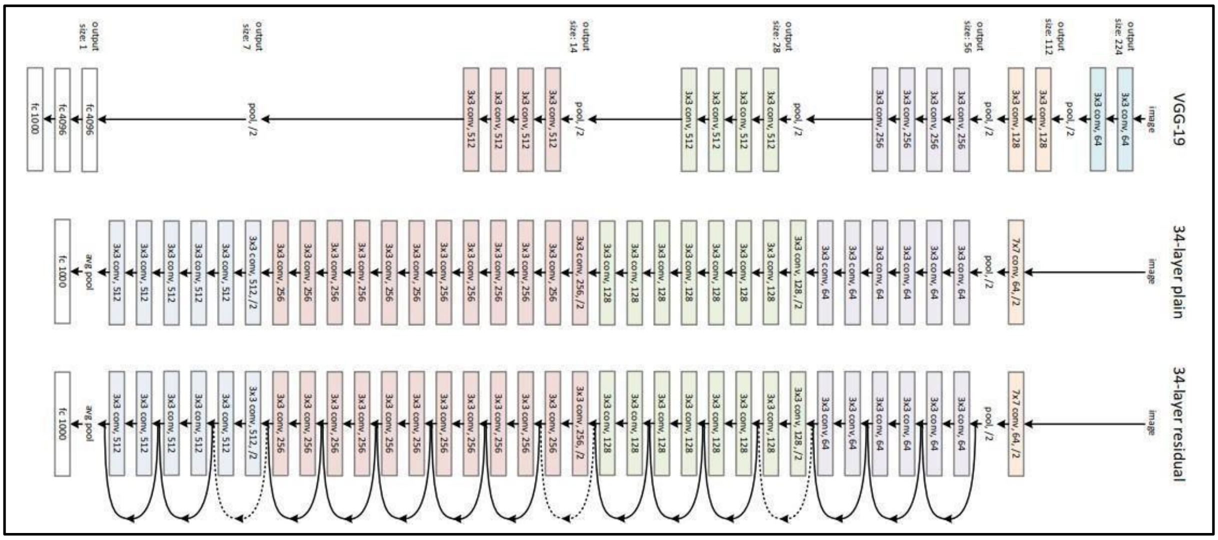


Skip (Shortcut) connection

The advantage of adding this type of skip connection is that if any layer hurts the performance of architecture then it will be skipped by regularization. So, this results in training a very deep neural network without the problems caused by vanishing/exploding gradients. The authors of the paper experimented on 100- 1000 layers of the CIFAR-10 dataset[1].

There is a similar approach called “highway networks”, these networks also use skip connection. Similar to LSTM these skip connections also use parametric gates. These gates determine how much information passes through the skip connection[2]. This architecture however has not provided accuracy better than ResNet architecture.

Network Architecture: This network uses a 34-layer plain network architecture inspired by VGG- 19 in which then the shortcut connection is added. These shortcut connections then convert the architecture into a residual network[3].



ResNet -34 architecture

Implementation: Using the Tensorflow and Keras API, we can design ResNet architecture (including Residual Blocks) from scratch. Below is the implementation of different ResNet architecture. For this implementation, we use the CIFAR-10 dataset. This dataset contains 60, 000 32×32 colour images in 10 different classes (aeroplanes, cars, birds, cats, deer, dogs, frogs, horses, ships, and trucks), etc. This dataset can be assessed from keras.datasets API function[4].

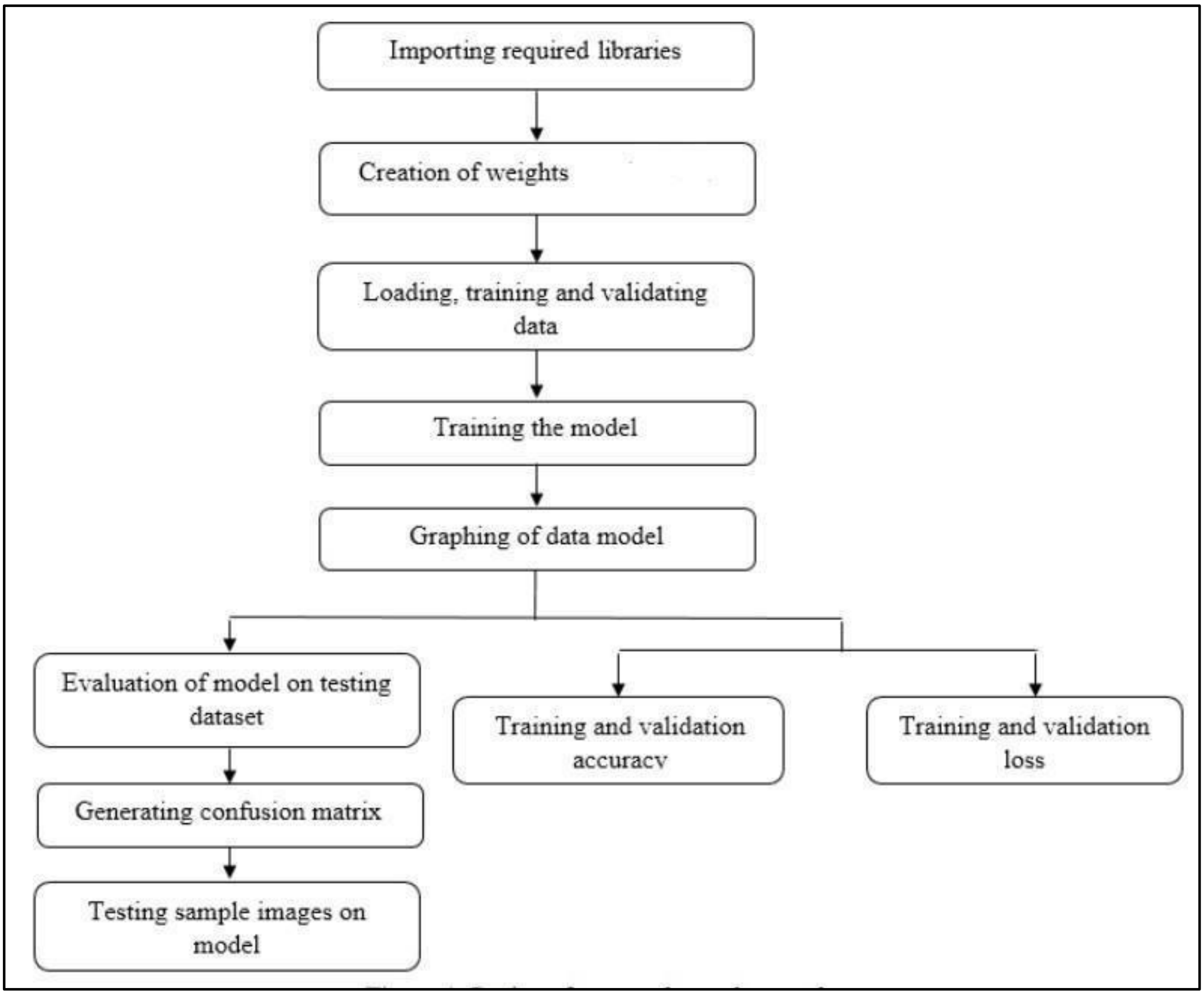
The system for animal species classification using Convolutional Neural Networks (CNN) with ResNet152V2 revolves around leveraging the robust capabilities of deep learning for accurate and efficient classification of animal species from images. At its core, the system utilizes the ResNet152V2 architecture, renowned for its depth and performance in image recognition tasks. Through a series of convolutional layers with residual connections, ResNet152V2 can extract hierarchical features from input images, capturing intricate patterns essential for discriminating between different animal species[5].

In this system input images of various animal species undergo preprocessing to ensure uniformity in size, resolution, and color channels. These images are then fed into the ResNet152V2 model, where they pass through multiple convolutional layers to extract high-level features. Following feature extraction, a classification layer maps the extracted features to specific animal species, generating probability distributions over the possible classes. Techniques such as data augmentation, dropout regularization, and transfer learning are employed to enhance the model's robustness and generalization capabilities[6]. By combining these elements, the system achieves accurate and efficient classification of animal species, facilitating diverse applications in wildlife monitoring, biodiversity conservation, and ecological research.

4. PROCESS FLOW

In the Process flow of the Animal Species Classification system using CNN with ResNet152V2, the process begins with the input of images containing various animal species. These images undergo preprocessing steps, including resizing, normalization, and standardization, to ensure uniformity and enhance the model's ability to extract relevant features. Subsequently, the preprocessed images are fed into the ResNet152V2 model, where they traverse through a series of convolutional layers with residual connections.

As the images pass through the convolutional layers, the ResNet152V2 model extracts hierarchical features at different levels of abstraction, capturing intricate patterns and distinguishing characteristics of different animal species[7]. These features are then propagated through the network, ultimately reaching the classification layer. Here, the extracted features are mapped to specific animal species using softmax activation, generating probability distributions over the possible classes. Finally, the system outputs the predicted animal species along with their corresponding probabilities, facilitating accurate classification and identification of animal species from input images. This data flow process enables seamless and efficient classification of animal species using CNN with ResNet152V2, contributing to various applications in wildlife conservation, ecological monitoring, and biodiversity research.



Process Flow Diagram for the Animal Species Classification

In the database design for Animal Species Classification using CNN with ResNet152V2, the primary focus is on storing and managing the image data required for training, validation, and testing of the classification model. The database structure typically consists of tables or collections to organize the image data, along with metadata such as labels indicating the corresponding animal species. Each entry in the database corresponds to an individual image, with fields to store the image file path or binary data, as well as associated labels or annotations.

Additionally, the database may incorporate tables or collections to store information related to model training and evaluation, such as hyperparameters, training/validation/test splits, and performance metrics[8]. This allows for comprehensive tracking and management of the experimental setup, facilitating reproducibility and scalability of the classification system. Overall, the database design plays a crucial role in supporting the data-intensive nature of the animal species classification task, providing a structured and efficient storage solution for image data and associated metadata throughout the model development lifecycle[10].

The for Animal Species Classification using CNN with ResNet152V2 illustrates the relationships between various entities involved in the system. The primary entities typically include "Images" and "Species." The "Images" entity represents individual images used for training and testing the CNN model. It includes attributes such as an image ID, file path, and metadata like image dimensions. On the other hand, the "Species" entity represents the different animal species to be classified. It contains attributes such as species ID and species name.

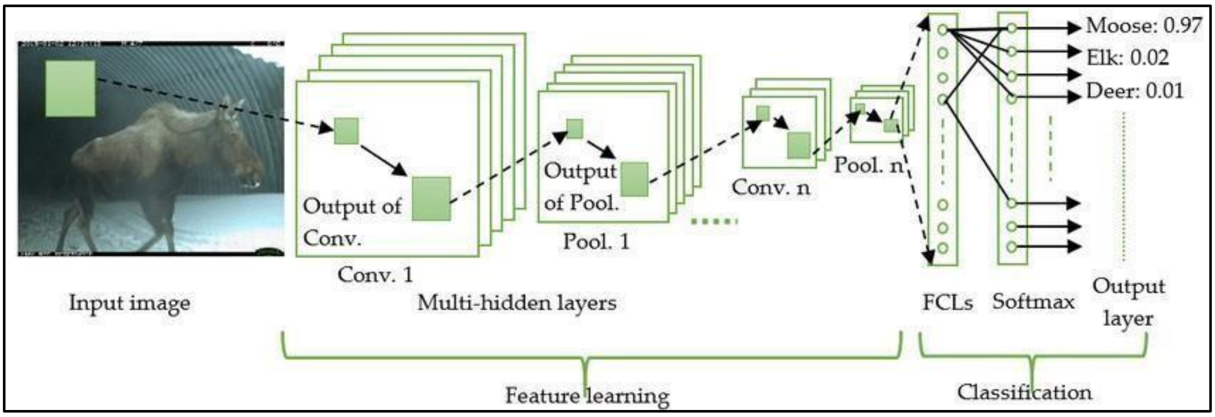


Fig 15. Feature Learning from the multi hidden layers

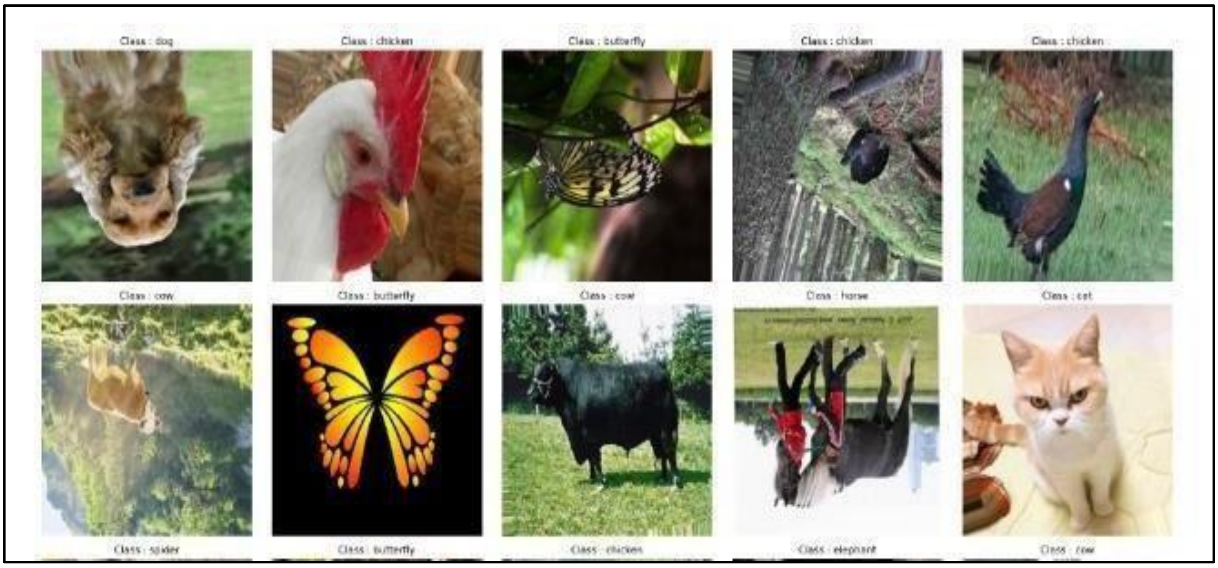
The relationship between these entities is typically represented as a one-to-many relationship, indicating that one species can have multiple images associated with it. Additionally, there may be auxiliary entities such as "Training Data" and "Validation Data," which store the relationships between images and their respective species for model training and evaluation[8]. Overall, the ER diagram provides a visual representation of the database schema, outlining the structure and relationships between entities essential for animal species classification using CNN with ResNet152V2.

5. RESULTS AND DISCUSSIONS

The input design for Animal Species Classification using CNN with ResNet152V2 involves the preparation and organization of image data to facilitate effective model training and evaluation. Firstly, the input design encompasses the collection of diverse image datasets containing various animal species, ensuring representation across different classes[4]. These images are then preprocessed to ensure consistency in size, resolution, and color channels, aligning them with the input requirements of the CNN model. Additionally, data augmentation techniques may be applied to increase the diversity of the training data and improve the model's ability to generalize to unseen samples[3]. Overall, the input design aims to provide a well-structured and standardized dataset that enables the CNN model to learn discriminative features for accurate classification of animal species[3].

Furthermore, the input design may involve the partitioning of the image dataset into training, validation, and testing sets to facilitate model development and evaluation. This partitioning ensures that the model is trained on a sufficiently large and diverse dataset while also validating its performance on unseen data. The training set is used to optimise the model's parameters through iterative optimization algorithms, while the validation set is employed to tune hyperparameters and assess the model's generalisation capabilities. Finally, thetesting set is utilised to evaluate the model's performance on unseen data and measure its classification accuracy. By carefully designing the input data pipeline, the CNN model can effectively learn the underlying patterns and characteristics of animal species, enabling accurate classification outcomes[5].

The result for Animal Species Classification using CNN with ResNet152V2 focuses on providing meaningful classification results that are interpretable and actionable. Once the CNN model is trained and deployed, the output design entails generating predictions for input images and presenting them in a comprehensible format. Typically, the output includes the predicted class label for each input image, indicating the animal species that the model believes the image belongs to. Additionally, the output may also include probabilities or confidence scores associated with each predicted class, providing insights into the model's confidence level for each classification. Moreover, the output design may incorporate visualization techniques to enhance the interpretability of the classification results[4]. This could involve displaying the input images alongside their predicted class labels and corresponding probabilities, allowing users to visually inspect the model's performance.



Identification and Classification based on CNN

Furthermore, the output may include metrics such as classification accuracy, precision, recall, and F1-score, providing a quantitative assessment of the model's performance. By designing the output to be informative and user-friendly,stakeholders can effectively interpret the classification results and make informed decisions based on the CNN model's predictions.

Exploring the Animal-10 Dataset: Obtaining Class Names and Counting Classes:

In this part of the project, we are setting the path to the Animal-10 dataset and obtaining a list of

class names from the dataset. We are then counting the number of classes and printing the original class names and the total number of classes[3]. This information is important for understanding the structure of the dataset and preparing it for use in a deep learning model.

```
Out[7]:
        {'cane': 4863,
         'cavallo': 2623,
         'elefante': 1446,
         'farfalla': 2112,
         'gallina': 3098,
         'gatto': 1668,
         'mucca': 1866,
         'pecora': 1820,
         'ragno': 4821,
         'scoiattolo': 1862}
```

The code prints the class names and the total number of classes in the Animal-10 dataset. Based on the output, we can observe that there are a total of 10 classes in the dataset and class names are in Italian.

These class names correspond to different types of animals, which are the targets of our image classification model. By knowing the number and names of the classes, we can better understand the structure of the dataset and prepare it for training the deep learning model.

Examining Class Distribution in the Animal-10 Dataset:

By obtaining the number of samples in each class, we can calculate the class distribution and determine whether the dataset is balanced or imbalanced.

This information can be used to determine whether any additional steps, such as data augmentation or class weighting, are necessary to balance the dataset and improve the performance of the deep learning model[4].

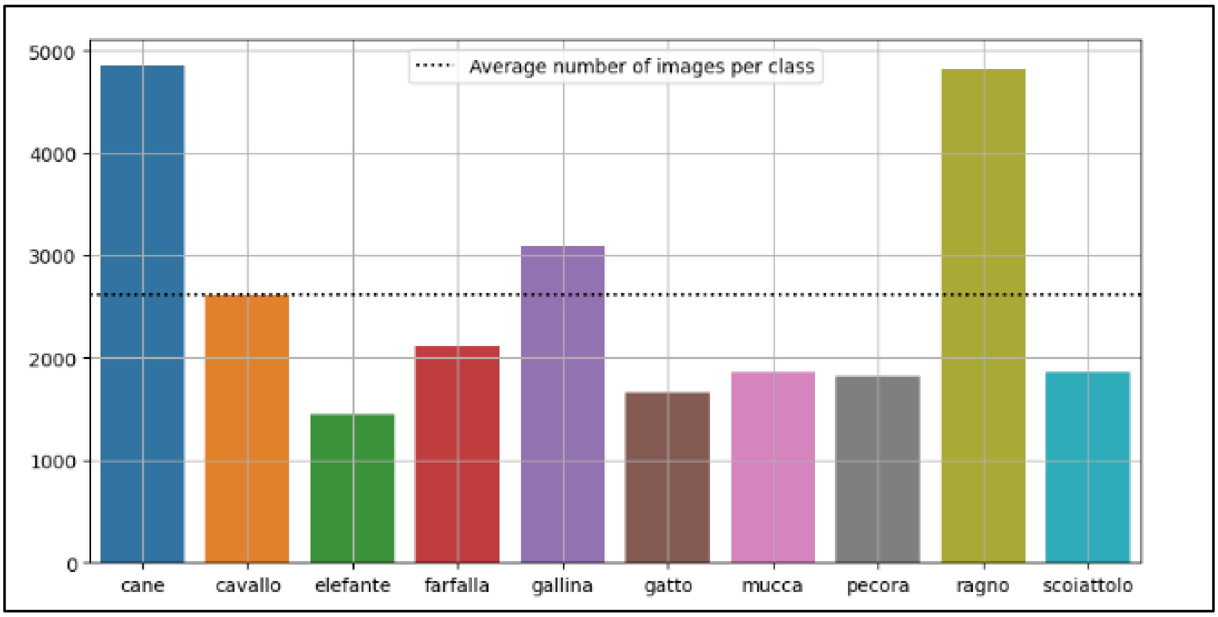
The code prints a list of integers, which correspond to the number of samples in each class of the Animal-10 dataset. Based on the output, we can observe that the class distribution is not balanced, as the number of samples in each class varies widely. The smallest class, "elefante", has only 1446 samples, while the largest class, "cane", has 4863 samples.

This imbalance could potentially affect the performance of a deep learning model trained on this dataset, as the model may be biased towards the majority class and have lower accuracy on the minority classes. Therefore, it may be necessary to use techniques such as data augmentation or class weighting to balance the dataset and improve the performance of the model.

Visualizing Class Distribution in the Animal-10 Dataset using a Pie Chart and Bar Graph:

The code below provides a visual representation of the class distribution of the Animal-10 dataset

using a pie chart, which can help us to better understand the dataset and identify any imbalances.



The bar graph further provides a visual representation of the distribution of the number of images in each class of the Animal-10 dataset, which can help us to better understand the dataset and identify any imbalances.

Evaluation of Pre-trained ResNet152V2 Model on Validation Dataset:

The line of code below evaluates the performance of the pretrained ResNet152V2 model on the validation dataset. The output of this line of code gives us an idea of how well the pre-trained model is performing on the validation data, which is a good indicator of its performance on unseen data. If the performance is good, it means that the pre-trained model has learned relevant features that can be useful for our classification task[9].

The output shows that the pre-trained model achieved a validation accuracy of 90.39% and a validation loss of 0.2987. This indicates that the model is performing quite well on the validation dataset.

	precision	recall	f1-score
Butterflies: 0	0.98	0.89	0.93
Chickens: 1	0.77	0.94	0.85
Elephants: 2	0.89	0.86	0.88
Horses: 3	0.99	0.91	0.94
Spiders: 4	0.93	0.95	0.94
Squirrels: 5	0.95	0.89	0.92
micro avg	0.93	0.91	0.92
macro avg	0.92	0.91	0.91
weighted avg	0.93	0.91	0.92
samples avg	0.91	0.91	0.91

Table 1 : Table Design

Here, we are visualising the performance of our pretrained ResNet152V2 model on the validation dataset. We randomly select an image and its corresponding label from the validation dataset, predict its label using the pre-trained model, and then plot the image along with its true label and predicted label. We repeat this process for 25 images and display the results using a 5x5 grid of subplots. This helps us to get an idea of how well our model is performing on the validation set and identify any

potential misclassifications.

From the output, we can see that the pretrained ResNet152V2 model is performing well on the validation dataset, as all the predictions made in the code are correct. This is a good indication that the model has learned to classify the images correctly and can be used for predicting the classes of unseen images.

6. CONCLUSION AND FUTURE ENHANCEMENT

Based on the results of the project, it can be concluded that training a ResNet152V2 model on the animal-10 dataset can achieve high accuracy in classifying different animals. The model was able to achieve an accuracy of 90.39% on the validation set, which is a good performance. Additionally, the results of the model predictions on sample images show that it is able to accurately classify animals with a high degree of confidence.

When comparing my work with the original project, it can be seen that both achieved similar results in terms of accuracy and model architecture. However, the original project used the entire dataset while I used only 10% of it. Overall, the project demonstrates the effectiveness of using pre- trained models like ResNet152V2 for image classification tasks, and highlights the importance of having a large and diverse dataset for achieving higher accuracy.

Through this study, a better neural network model can be obtained to realise the classification of animal images, and further research of neural networks is realised by combining the machine learning

and deep learning technology with the ecological protection.

REFERENCES

[1] Hanguen Kim, Jungmo Koo, Donghoonkim, Sungwoo Jung, Jae-Uk Shin, Serin Lee, Hyun Myung, “Image-Based Monitoring of Jellyfish Using Deep Learning Architecture”, IEEE sensors journal, vol. 16, no. 8, 2016.

[2] Carlos Silva, Daniel Welfer, Francisco Paulo Gioda, Claudia Dornelles,” Cattle Brand Recognition using Convolutional Neural Network and Support Vector Machines ”, IEEE Latin America Transactions, vol. 15, no. 2, pp. 310- 316, 2017.

[3] S. Yang, L. Bo, J. Wang, and L. G. Shapiro, "Unsupervised template learning for finegrained object recognition," in Advances in Neural Information Processing Systems, 2012, pp. 3122-3130. [4]. Dhruv Rathi, Sushant Jain, Dr. S. Indu, “Underwater Fish Species Classification using Convolutional Neural Network and Deep Learning”, International Conference of Advances in Pattern Recognition, 2017.

[4] Mohamad Aqib Haqmi Abas, Nurlaila Ismail, Ahmad Ishan Mohd Yassin, Mohd Nasir Taib, “VGG16 for plant image classification with transfer learning and data augmentation”, International Journal of Engineering & Technology, 2018.

[5] Kaggle.com,Animals10dataset[Online],Availableat : https://www.kaggle.com/alessiocorrado99/animals10.

[6] M. A. Al-antari, P. Rivera, M. A. Al-masni, E. Valarezo, G. Gi, T. Y. Kim,

[7] H. M. Park, and T. S. Kim, "An automatic recognition of multiclass skin lesions via Deep Learning Convolutional Neural Networks," In Conference: ISIC2018: Skin Image Analysis Workshop and Challenge. 2018.

[8] L. Hu, and Q. Ge, "Automatic facial expression recognition based on MobileNetV2 in Real- time," In Journal of Physics: Conference Series, vol. 1549, no. 2, 2020.

[9] Hussain M, Bird JJ, Faria DR. A study on cnn transfer learning for image classification. InUK Workshop on computational Intelligence 2018 Sep 5 (pp. 191-202). Springer, Cham.

[10] Gao, Y., Mosalam, K.: Deep transfer learning for image-based structural damage recognition. Comput. Aided Civ. Infrastruct. Eng. (2018).