Bird Detection using Siamese Neural Network

Rushil Gupta, Shashank Pandey, D Vanusha

Abstract: The proposed system uses deep neural networks for identifying bird species. The model will be trained on bird images that are coming in the endangered species category. The application can also handle new data points, unlike existing systems that require model re-training for accommodating new data.

The system can identify bird species in a large view of the image. The model will be trained using a convolutional neural network-based architecture called Siamese Network. This network is also called one-shot learning which means that it requires only few training example for each class.

Existing models use image processing techniques or vanilla convolutional neural networks for classifying bird images. These models cannot accommodate new images and have to be retrained to do so. There is no commercially available system that can detect a species of bird in high resolution / large image. While in the Siamese network we only have to add new data, there is no need to retraining the neural network.

Keywords: Neural networks, Model deployment, Edge AI, Deep learning

I. INTRODUCTION

Now a days, In case of object detection we commonly use convolutional neural network. The general process of a neural network is to take image as input and apply some filter and pooling to extract only important features of the image. We can apply filter more than one time like the first filter will focus on background color or shape. Second filter will focus mainly on image. in this way it will filter the important features on image. In the end we create a flatten layer that contains the important features related to image. After this we can apply some weights and bias to flatten the layer. We have to apply activation functions to increase the efficiency. In this way it consists of a lot of layers and in the last layer we apply mostly sigmoid or tanh as activation function. The key important role of this model is to use loss function and learning rate. which will make model to learn. For each forward process we have to apply one backward process which will use these loss function and learning rate to learn the model and change the value of weighs and bias.

In this project we are using some latest deep learning architecture to increase the accuracy of the model. Our project consists of Siamese neural networks.

Revised Manuscript Received on May 30, 2020.
* Correspondence Author
 Rushil Gupta*, SRM Institute of Science and Technology rg1802@srmist.edu.in
 Shashank Pandey, SRM Institute of Science and Technology sp5935@srmist.edu.in
 D Vanusha, SRM Institute of Science and Technology vanushad@srmist.edu.in

In vanilla convolutional neural networks the efficiency of a model is directly proportional to size of data-set and quality of data-set. So sometimes when large data-set is not present we apply some algorithm like augmentation to increase the size of the data-set which sometimes add noises. Which reduces the efficiency. It demands a large data-set which is very difficult in some cases. Second, if we test our model for classes which are not present in our data-set it will not detect. now to add that class we have to first add these classes to our data-set the re-train the whole model again.

We have to apply previous process again. Which now require more computational power and sometimes it need to change the values of some hyper-parameters.

In some application we do not have data in large amount and we have to set the classes dynamically as in case of bird detection we have to increase the classes as we face with any new specie of bird. In that cases training new model is very difficult.

On the other hand, in a one shot classification or Siamese neural network, we have to use only one image of each specie and when we have to add new species, it does not require to train the whole model again it will only add the features of that image.

II. RELATED WORK

2.1 Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial Vehicle Imagery

In this article, wild birds are monitored with the objectives of identifying their habitat and making an estimate of their population. The data set used here contains aerial photographs with diverse images of various birds habitats and on farmland. The models like Faster Region-based Convolutional Neural Network (R-CNN), Single Shot Multi Box Detector (SSD), You Only Look Once (YOLO) , Region-based Fully Convolutional Network (R-FCN), and Retinanet were made and the overall tradeoff of performance was measured. The results from the experimentation were in favor of Faster R-CNN to be most accurate model.

2.2 Domestic Cat Sound Classification Using Learned Features from Deep Neural Nets

The area this paper deals with the classification of cat sounds using Machine Learning. Using Machine learning the size of the data can be small for the labeled class. The comparison for five machine learning algorithms was made with respect to data increased while augmenting, the features learned from Pre-Trained CNN or CDBN (unsupervised).

2.3 GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification

The vast breakthrough of deep learning methods particularly CNN in a broader range of computer vision tasks but generating data sets or obtaining them is quite challenging therefore in this article the methods of generating medical images were covered using Generative Adversarial Networks (GANs).

CNN was trained using classic data augmentation and the performance was compared with synthetic data augmentation.

2.4 Environment Sound Classification Using a Two-Stream CNN Based on Decision-Level Fusion

With the increase in the number of researchers, the methods to classify the environment sound has been exploited. The paper proposes the two combined features to provide a detailed summary of environment sounds. After that a four-four layer CNN was used to catalyze the performance of ESC with the proposed combined features.

2.5 Interpreting and Explaining Deep Neural Networks for Classification of Audio Signals

This paper traverse that how illustratable the neural networks are when comes to the audio domain. The previous researches include layer-wise relevance propagation (LRP). The method used here is LRP to identify the useful features for 2 neural network architecture that can process either spectrogram or waveform. The results are in favor of concluding that networks are well reliant on features as by LRP.

2.6 End-to-End Environmental Sound Classification using a 1D Convolutional Neural Network

This article deals with the signals in the form of audio waves of variable length and as the signal splits to overlapped frames are formed using a sliding window. The different architectures are considered for different input sizes and the performance has been measured. the proposed approach has few parameters contrasted with different models found in the writing, which diminishes the measure of information required for training.

2.7 Bird Species Identification using Deep Learning

The article focuses on utilizing the techniques and methods from deep learning to classify the bird species. The algorithm used here is DCNN and the image is converted to greyscale using tensor flow. The various nodes are compared with test and train dataset the prediction score is high.

2.8 Birds Voice Classification using ResNet

The paper deals with the classification and monitoring of birds using their sound. The deep learning techniques and literature surveys were done on traditional voice recognition techniques and the issues in various methods were noted and resolved using the proposed system architecture.

2.9 Deep Learning Case Study for Automatic Bird Identification.

A customized winged creature conspicuous evidence system is used to study offshore wind develops in many places. Unquestionably, a radar is used mainly to recognize flying winged creatures, yet outside information is required for genuine ID. We had used visual camera pictures as outside data. The proposed structure for modified fowl conspicuous confirmation contains a radar, a motorized video head and a singular point of convergence reflex camera with a zooming point of convergence. A convolutional neural framework arranged with a significant learning figuring is applied to the image game plan. We moreover propose a data extension method in which pictures are turned and changed over according to the perfect concealing temperatures. The keep going unmistakable verification relies upon a blend of parameters gave by the radar and the gauges of the image classifier. The affectability of this proposed structure, on a dataset containing 9312 genuinely taken novel pictures achieving 2.44×10^6 extended enlightening record, is 0.9463 as an image classifier. The zone under recipient working trademark twist for two key fowl species is 0.9993 (the White-followed Eagle) additionally, 0.9496 (The Lesser Black-upheld Gull), exclusively. We proposed a novel system for modified winged creature separating confirmation as a genuine application. We demonstrated that our data extension system is fitting for picture game plan issue and it basically extends the introduction of the classifier.

III. ALGORITHMS USED

3.1 Convolutional Neural Networks

In neural systems, Convolutional neural system (ConvNets or CNNs) is one of the primary classifications to do image detection, pictures classificstion. Items identifications, acknowledgment faces and so on., are a portion of the zones where CNNs are generally utilized.

CNNs give an ideal design for image detection and image classification. By doing parallel computation and using GPUs, CNNs are a key innovation hidden new improvements algorithm to perform experiments on images.

For instance, in medical fields the CNNs is used to look at huge numbers oh reports which are consists of images to predict the new one.. CNNs likewise empower self-driving autos to recognize items and figure out how to differentiate between a road sign and a passerby.

A convolutional neural system can have many layers that each figure out how to identify various highlights of a image. Channels are applied to each preparation picture at various goals, and the yield of each convolved picture is utilized as the contribution to the following layer. The channels can begin as straightforward highlights, for example, splendor and edges, and increment in intricacy to highlights that remarkably characterize the article.

Fig 1: Convolutional Neural Network
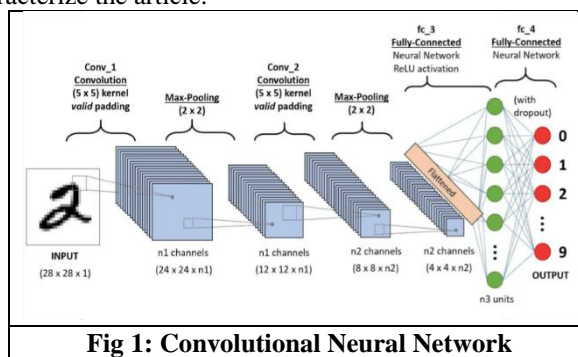
3.2 Siamese Neural Network

Siamese neural system is a class of neural system designs that contain at least two indistinguishable subnetworks. indistinguishable here means they have a similar arrangement with similar parameters and loads. Parameter refreshing is reflected crosswise over both subnetworks.

Siamese NNs are mainstream among errands that include discovering similitude or a connection between two practically identical things. A few models are rework scoring, where the sources of info are two sentences and the yield is a score of how comparative they are; or signature check, where make sense of whether two marks are from a similar individual. For the most part, in such undertakings, two indistinguishable subnetworks are utilized to process the two sources of info, and another module will take their yields and produce the last yield. The image beneath is from Bromley et al (1993). They proposed a Siamese engineering for the mark check task.
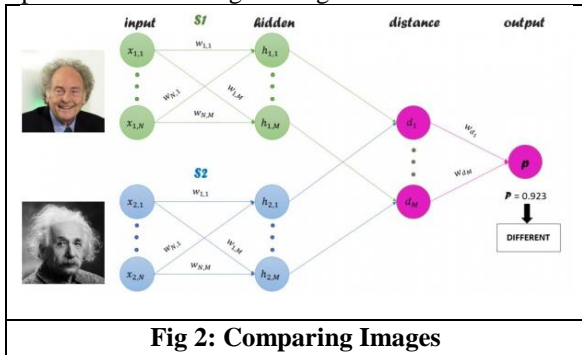

Fig 2: Comparing Images

IV. PROPOSED SYSTEM

4.1 Bird Species Detection

A siamese neural network algorithm is applied on some birds species. We will take two images, both the images are feed to convolutional neural network. We will apply filters and max pooling to extract the features of that particular image. We stop convolutional neural network till we get the flatten layer. Which contains the more precised features of that image. Now we will take the flatten vector of these two images and calculate the distance between these two vectors. The value then pass through the activation fuction like sigmoid or tanh. We will calculate the probability of distance between these two images. If the probability is near to 1 then they are same otherwise they are different.
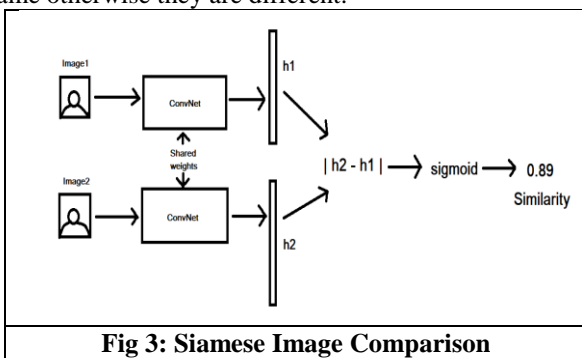

Fig 3: Siamese Image Comparison

4.2 Training the model

In the siamese neural network, the training is preceded by taking three images.

First is the anchor image, second one is a positive image and third one is a negative image. This special architecture help us reduce the dataset size. Now we can train our model by creating the flatten vector. Flatten vector is made by applying filters and max pooling on images. Now we will compare the anchor image with positive and negative image. This is done by applying the loss function. Which will apply cosine rule one between anchor image and positive image and one between anchor image and negative image.

$$Loss = \sum_{i=1}^{N} \left[\|f_i^a - f_i^p\|_2^2 - \|f_i^a - f_i^n\|_2^2 + \alpha \right]_+$$
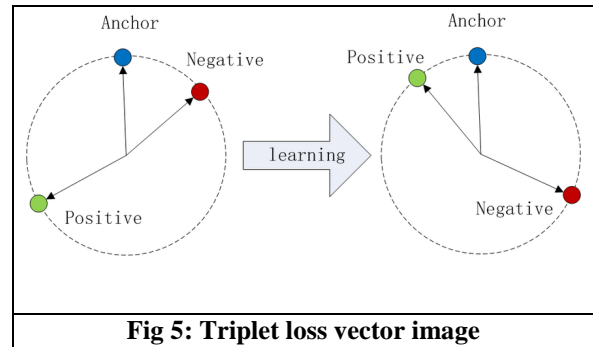
Fig 4: Siamese loss equation

Fig 5: Triplet loss vector image

General Explanation:

$$loss = max(d(a,p)-d(a,n) + beta,0)$$

The above formula shows the triplet loss function. Here a represents the flatten vector of anchor image. p represents the flatten vector of positive image and n represents the flatten vector of negativer image. Another parameter which is beta which is hyperparameter. which defines the range of dissimilarities. By calculating the loss function we can calculate the gradients by using learning rate and by using gradients we can update the weights and biases.

V. RESULT

The following are the result of the experiments :
- Training on 6 species -
  - Bengal Florican
  - Great Hornbill
  - Great Indian Bustard
  - Red Headed Vulture
  - Spoon Billed Sandpiper
  - Spot-billed pelican

We had taken 30 images for each species.

Accuracy on 180(30*6) images when trained using the siamese network = 88%.

VI. CONCLUSION

This paper manages the programmed fowl species identification from feathered creature pictures. We present a progression of tests led in a dataset created by more than 6,000 images from 200 diverse winged creature species. In this work, we found that without PC vision we had the option to characterize winged creatures into expansive species classifications very well, however less so within explicit species classifications because of likenesses in qualities of firmly related feathered creature species. We additionally found that utilizing PC vision on HOG+RGB highlights of the general winged animal gives preferable outcomes over HOG

alone, which bodes well since there is more data and in light of the fact that the shades of the flying creature are basic in recognizing their species. Including feathered creature head and nose parts as highlights just gave slight gains in exactness, perhaps due di↵erences in frontal versus side perspectives on the winged animal. Generally speaking, move learning of the pre-prepared AlexNet CNN has the most encouraging outcomes. We found that calibrating was delayed because of numerous layers of backpropagation and that utilizing fixed elements gave the most sensible outcomes as far as computational cost, precision, and forestalling over-fitting.

ACKNOWLEDGEMENT

REFERENCES

1. Suk-Ju Hong, Yunhyeok Han, Sang-Yeon Kim, AhYeong Lee, Ghiseok Kim , "Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial
Vehicle Imagery", Sensors 2019, 19(7), 1651
2. Yagya Raj Pandeya, Dongwhoon Kim and Joonwhoan Lee. "Domestic Cat Sound Classification Using Learned Features from Deep Neural Nets" Appl. Sci. 2018, 8(10), 1949;
3. Maayan Frid-Adar, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, Hayit Greenspan. "GAN-based synthetic medical image augmentation for increased
4. Yu Su, Ke Zhang, Jingyu Wang, Kurosh Madani. "Environment Sound Classification Using a Two-Stream CNN Based on Decision-Level Fusion" Sensors 2019, 19(7),1733
5. Sören Becker, Marcel Ackermann, Sebastian Lapuschkin, Klaus-Robert Müller, Wojciech Samek. "Interpreting and Explaining Deep Neural Networks for Classification of Audio Signals" CoRR 2018
6. Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, Quoc V Le. "AutoAugment: Learning Augmentation Policies from Data". CoRR 2019
7. David E Rumelhart, Geoffrey E Hinton, Ronald J Williams. "Learning representations by backpropagating errors" Nature (1986)
8. Pralhad Gavali, Prachi Abhijeet Mhetre, Neha Chandrakhant Patil, Nikita Suresh Bamane, Harshal Dipak Buva. "Bird Species Identification using Deep Learning" IJERT 2019
9. Aarti Madhavi, Rajni Pamnani. "Deep Learning Based Audio Classifier for Bird Species" IJSDR 2018
10. Juha Niemi, Juha T Tanttu. "Deep Learning Case Study for Automatic Bird Identification". Appl. Sci. 2018
11. Tejashri Manohar Mhatre, Srijita Bhattacharjee. "Birds Voice Classification using ResNet" IJEDR 2018
12. Nitish Shrivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting" Journal of Machine Learning Research 2014
13. Sajjad Abdoli, Patrick Cardinal, Alessandro Lameiras Koerich. "End-to-End Environmental Sound Classification using a 1D Convolutional Neural Network" CoRR 2019.

Retrieval Number: E2468039520/2020©BEIESP
DOI: 10.35940/ijitee.E2468.059720
Journal Website: www.ijitee.org

1171

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication