Bird Species Identification Using Modified Deep Learning (CNN) Model by Integrating Cascaded Softmax Layer

K. AMARAVATHI, M.Tech, (Ph.D) [#1], N. LOKESWARI, M.Tech, (Ph.D.) [#2]

[#1] Assistant Professor, [#2] Assistant Professor,

Department of Computer Science & Engineering,

Anil Neerukonda Institute of Technology & Sciences,

Sanghivalasa, Visakhapatnam, AP 531162.

ABSTRACT

Now a days for identifying or predict any living beings, we should have proper idea about that living beings. For example if we want to predict some animals,plants,birds,dogs and so on, we need to know clearly about that living beings, before we try to predict the end result. In general we try to take images of those living beings especially birds and then try to classify the birds based on its features and then try to decide the species name. In this proposed application we develop a fine-grained image classifier using a general deep convolutional neural network (D-CNN). In order to increase the accuracy of our proposed model we try to modify the DCNN model from the following two aspects. First, to better model the h-level hierarchical label structure of the fine-grained image classes contained in the given training data set, we introduce h fully connected (fc) layers to replace the top fc layer of a given DCNN model and train them with the cascaded softmax loss. Second, we propose a novel loss function, namely, generalized large-margin (GLM) loss, to make the given DCNN model explicitly explore the hierarchical label structure and the similarity regularities of the fine-grained image classes. The GLM loss explicitly not only reduces between-class similarity and within-class variance of the learned features by DCNN models but also makes the subclasses belonging to the same coarse class be more similar to each other than those belonging to different coarse classes in the feature space. By implementing these two aspects on current DCNN model, we try to propose a new modified DCCN model known as MDCNN model, which can increase the accuracy of deep learning models in order to predict the bird species very accurately. By conducting various experiments on our proposed model by comparing with several pre-trained models such as AlexNet, GoogLeNet, and VGG, using three benchmark data sets.Our simulation results clearly demonstrate the accuracy and efficiency of our modified MDCNN model is standing on top of all primitive models.

Keywords:

Deep Convolutional Neural Network, Cascaded Softmax Loss**,** Generalized Large-Margin (GLM) Loss**,** Fully Connected.

1. INTRODUCTION

Fine-Grained image classification aims to recognize subordinate classes of some base class, such as different models of cars [1]–[5], species of birds [5]–[9], variants of aircrafts [10], [11], and so on. It has a wide range of applications, such as vehicle model recognition for video surveillance, fine-grained image content annotation, vertical search, and so on. The challenges of fine-grained image classification mainly come from the following two aspects: between-class similarity and within-class variance [12]–[16].

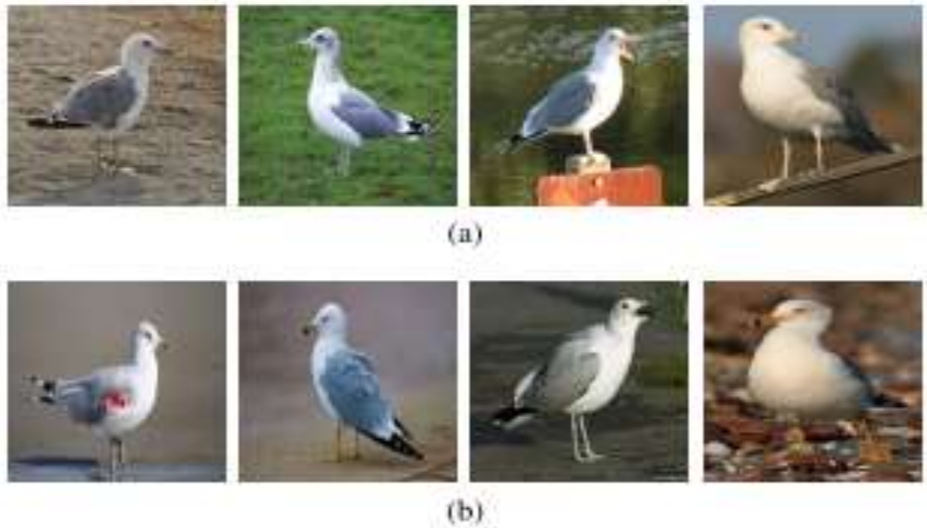


Fig. 1. Sample images of (a) California gull and (b) ringed-beak gull. The salient difference between them lies in the patterns on their beaks.

On one hand, visual differences between different fine-grained classes could be very small and subtle. On the other hand, instances belonging to the same fine-grained class may have significantly varied appearances due to different locations, viewpoints, poses, lighting conditions, and so on. For example, the “California gulls” shown in Fig. 1(a) are visually quite similar to the “ringed-beak gulls” in Fig. 1(b), and the only salient difference between them lies in the patterns on their beaks. Meanwhile, caused by differences in pose, viewpoint, and lighting condition, the “California gulls” in the different images (so as the “ringed-beak gulls”) exhibit remarkably different appearances between each other. From the above figure 1,we can clearly see that all the birds are almost similar and we cannot easily judge the species of that bird by viewing the bird manually,but if we come across the BEAK, they have small minor difference from one bird to another,which is the only way to identify the birds based on species.

2. LITERATURE SURVEY

Literature survey is that the most vital step in the software development process. Before developing the new application or model, it's necessary to work out the time factor, economy, and company strength. Once all these factors are confirmed and got approval then we can start building the application. The literature survey is one that mainly deals with all the previous work which is done by several users and what are the advantages and limitations of those previous models. This literature survey is mainly used for identifying the list of resources to construct this proposed application.

MOTIVATION

1) Bird Species Classification Using Deep Learning Approach

AUTHORS: SHRIHARSHA and TUSHARA

In this paper, the authors mainly concentrated on the classification of birds and how the species can be identified from bird images. Birds are the warm-blooded vertebrates constituting of class Aves, almost there are more than ten thousand living species present in the world and different appearances. In order to classify those many of birds in manual way is very tedious job where the humans require a lot of knowledge about all the birds and its species, as it requires lot of expertise in the field of Ornithology. In this paper the authors presented a new model in CNN where the bird species can be identified based on several bird images which are collected from google. The proposed model was mainly tested and trained on more than 9000 images which are gathered from twenty different bird species and the check the CNN model to find out the accuracy. By conducting various experiments on our proposed model with current dataset, we finally came to a conclusion that our model can achieve the accuracy of 98% when tested with the test datasets.

2) Bird Species Identification using Deep Learning on GPU platform

AUTHORS: J.Saira Banu

In this paper, the authors mainly discussed about deep Convolutional Neural Networks for bird species identification. In this paper, the author mainly present new probabilistic models for identifying bird species from audio recordings. He introduced a new independent syllable model and try to consider the new model into two main categories which can be used to aggregate the frame level features within a syllable. In this article the author described that use of CNN model is very efficient for bird species identification based on GPU and this can increase the chance of accuracy.

3) Image Bird Species Identification.

AUTHORS: Bipin Kumar Rai

In this paper, the author mainly explained about the process of bird species identification based on images. In general a lot of people will have a hobby to identify the different birds and its species by taking assistance from several text books or materials. To provide birdwatchers a handy tool to admire the beauty of birds, the authors developed a new model which can able to assist any new users who don't have proper knowledge about all the birds in in recognizing species of birds using a software based on the concept of image recognition. This developed model can act like a software for identifying the image by comparing the model with a trained model and then predict the bird species.

3. EXISTING SYSTEM AND ITS LIMITATIONS

In the existing system, there was no concept like bird species prediction using CNN models. All the prediction is done using manual approach or by using primitive Machine Learning models. In the ML we can able to classify whether bird is identified or not and what is the predicted species name for that bird,but those models cannot classify the records with accuracy and parameters.

LIMITATIONS OF THE EXISTING SYSTEM

1. All the existing schemes are limited to the few classes classification only.
2. All the existing systems are failed to classify the bird images by having manual method or ML algorithms.
3. All the existing ML approaches try to classify the birds data if the dataset is small.
4. All the existing ML approaches fail to classify the birds and its species if the data set is very large.
5. There is no accurate model to classify the birds and its species for detecting and prediction of accuracy of that image.

4. PROPOSED SYSTEM AND ITS ADVANTAGES

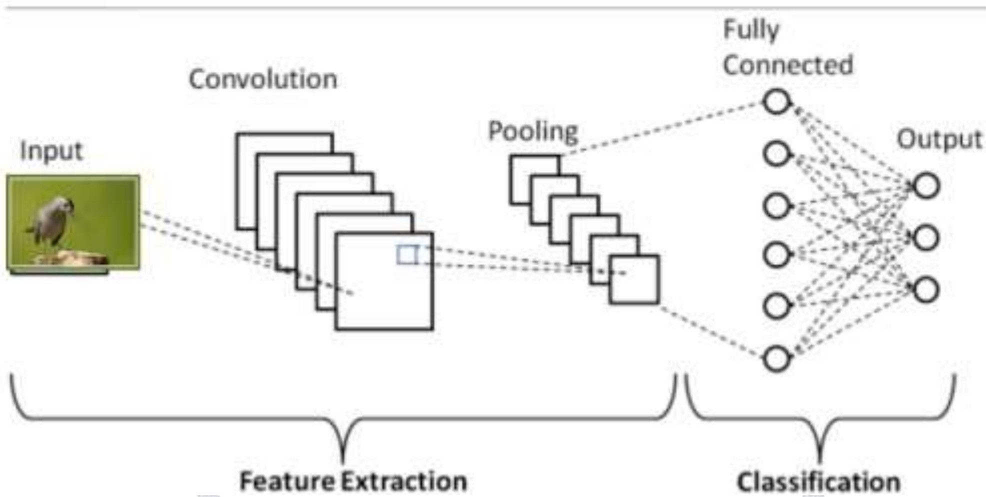
In this proposed application we develop a fine-grained image classifier using a general deep convolutional neural network (D-CNN). In order to increase the accuracy of our proposed model we try to modify the DCNN model from the following two aspects. First, to better model the h-level hierarchical label structure of the fine-grained image classes contained in the given training data set, we introduce h fully connected (fc) layers to replace the top fc layer of a given DCNN model and train them with the cascaded softmax loss. Second, we propose a novel loss function, namely, generalized large-margin (GLM) loss, to make the given DCNN model explicitly explore the hierarchical label structure and the similarity regularities of the fine-grained image classes. The GLM loss explicitly not only reduces between-class similarity and within-class variance of the learned features by DCNN models but also makes the subclasses belonging to the same coarse class be more similar to each other than those belonging to different coarse classes in the feature space. By implementing these two aspects on current DCNN model, we try to propose a new modified DCCN model known as MDCNN model, which can increase the accuracy of deep learning models in order to predict the bird species very accurately.

ADVANTAGES OF THE PROPOSED SYSTEM

1. The proposed scheme is very accurate in identification of birds and its species.
2. The proposed system gives accurate recommendation for the Ornithologist for classifying new bird species .
3. The proposed system is capable of classifying a large set of bird species very accurately.
4. The proposed system is advancement of primitive CNN model and hence it is termed as MDCNN.

5. PROPOSED MDCNN MODEL FOR BIRD SPECIES IDENTIFICATION

In this section we try to discuss about proposed modified deep convolution neural network (MDCNN) model which is used to detect bird species accurately and efficiently.



The Application is mainly divided into 4 modules. They are as follows:

1. Convolution Layer
2. Rectified Linear Unit (RELU) Layer
3. Pooling Layer
4. Fully Connected layer

A) CONVOLUTION LAYER

A convolution is defined as an operation on two functions. In image analysis, one function consists of input values (e.g. pixel values) at a position in the image, and the second function is a filter (or kernel) each can be represented as array of numbers. Computing the dot product between the two functions gives an output. The filter is then shifted to the next position in the image as defined by the stride length. The computation is repeated until the entire image is covered, producing a feature (or activation) map.

B) RECTIFIED LINEAR UNIT (RELU) LAYER: The RELU layer is an activation function that sets negative input values to zero. This simplifies and accelerates calculations and training, and helps to avoid the vanishing gradient problem. Mathematically it is defined as:

$$f(x)=\max(0,x).$$

Where x is the input to the neuron. Other activation functions include the sigmoid, tanh, leaky RELUs, Randomized RELUs and parametric RELUs.

C) POOLING LAYER

The Pooling layer is inserted between the Convolution and RELU layers to reduce the number of parameters to be calculated, as well as the size of the image (width and height, but not depth). Max-pooling is most commonly used; other pooling layers include Average pooling and L2-normalization pooling. Max-pooling simply takes the largest input value within a filter and discards the other values; effectively it summarizes the strongest activations over a neighborhood. The rationale is that the relative location of a strongly activated feature to another is more important than its exact location.

D) FULLY CONNECTED LAYER

The final layer in a CNN is the Fully Connected Layer, meaning that every neuron in the preceding layer is connected to every neuron in the Fully Connected Layer. Like the convolution, RELU and pooling layers, there can be 1 or more fully connected layers depending on the level of feature abstraction desired. This layer takes the output from the preceding layer (Convolutional, RELU or Pooling) as its input, and computes a probability score for classification into the different available classes

Once the fully connected layers process each and every input data ,then corresponding data is matched with its previous records and finally classified the bird and its species.Here this model uses Softmax classifier which can greatly increase the accuracy of prediction compared with primitive CNN models in order to classify the birds and its species very accurately.

DCNN MODIFICATION AND CASCADED

Softmax Loss For the fine-grained image classification problem with h-levels of class labels, we modify the given DCNN model by replacing its top fc layer with the h fc layers and train it using the cascaded softmax loss function. To ease the explanation, and without loss of generality, we describe the use of AlexNet to classify an image data set with two levels of class labels. DCNN modifications for other fine-grained image classification problems can be derived by analogy

6. EXPERIMENTAL RESULTS

Implementation is a stage where the theoretical design is converted into a programmatic manner. In this proposed application we try to use PYTHON as a programming language in which Google Collaboratory or Jupiter Notebook as a working platform to process the current application.

STEP 1: IMPORTING ALL NECCESARY LIBRARIES

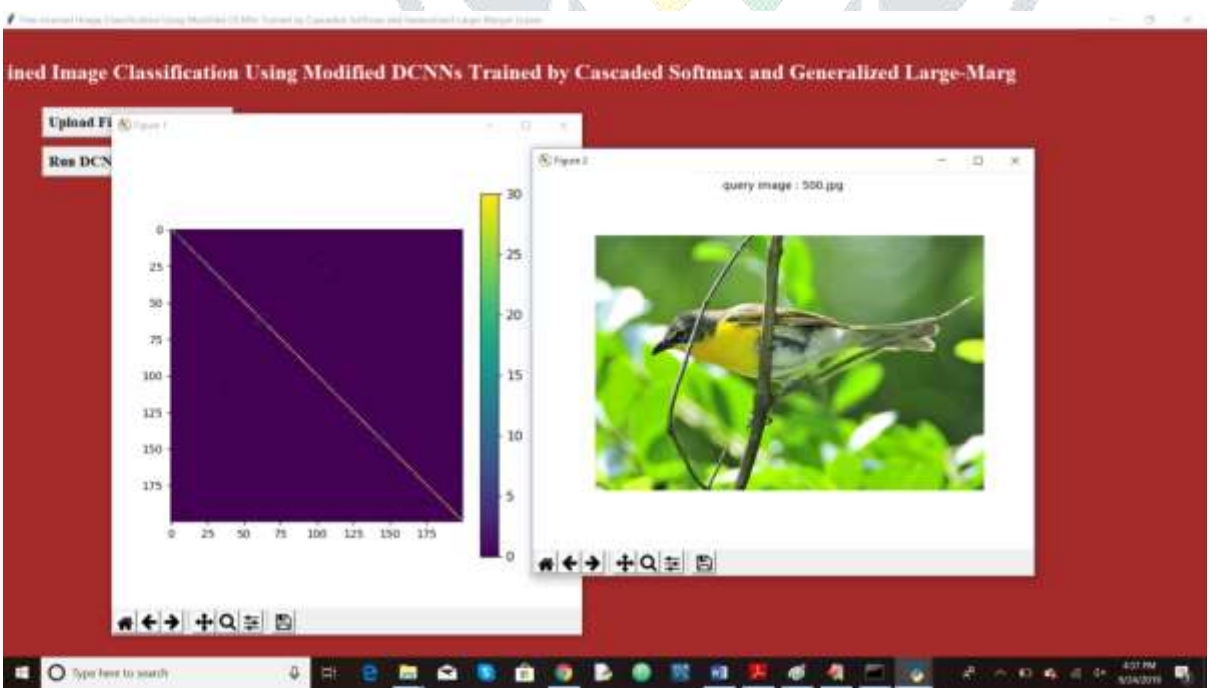
```
pip install sklearn
pip install scikit-image
```

STEP 2: double click on **run.bat** file to get below screen

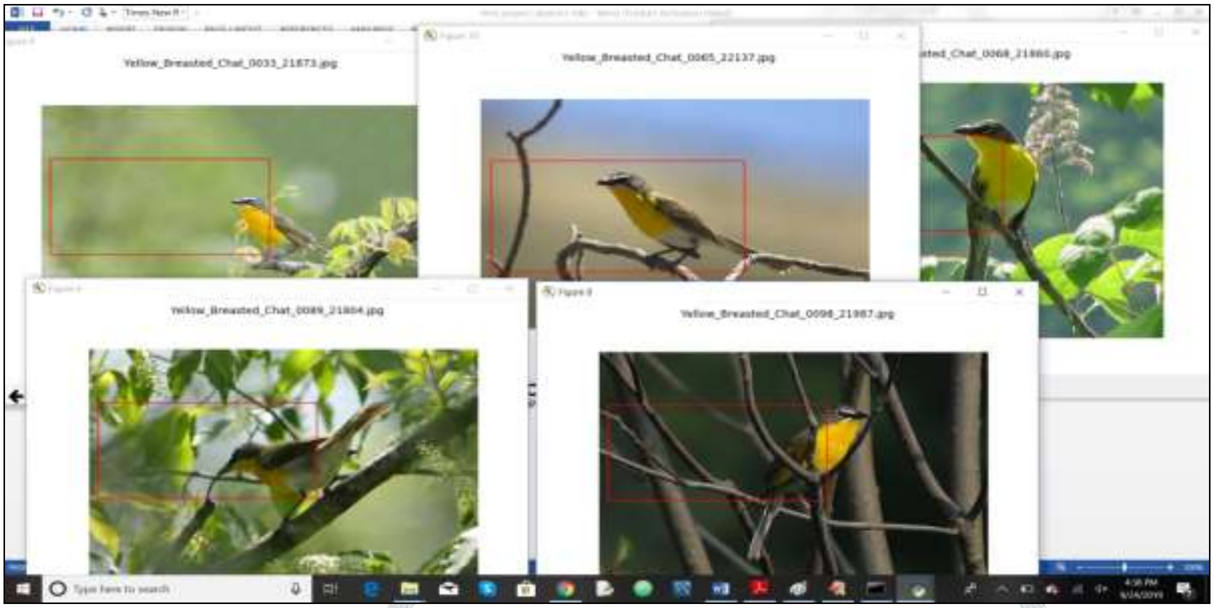
STEP 3: UPLOAD FINE GRAINED IMAGES

In above screen I am selected one image and now click on open button to get below screen

STEP 4: RUN DCNN Model



STEP 5 : Start Testing DCNN with Train Images .



In above screen we can see we got five images as search result as all images are different but they got searched based on body part color similarity. Their matched part are highlighted with red colour rectangle. With normal eyes they look different but by using this technique we can distinguish or identify them.

7. CONCLUSION

In this current work we for the first time designed and implemented an application using modified CNN model in order to find the bird species accurately and efficiently. A novel DCNN-based framework is proposed to improve fine-grained image classification accuracy. We improve the fine-grained image classification accuracy of a DCNN model from the following two aspects. First, we introduce h fc layers to replace the top fc layer of a given DCNN model and train them with the cascaded softmax loss to better model the h-level hierarchical label structure of the fine-grained image classes. Second, we propose the GLM loss to make the given DCNN model explicitly explore the hierarchical label structure and the similarity regularities of the fine-grained image classes. The proposed fine-grained image classification framework is independent of the DCNN structure. Comprehensive experimental evaluations of several general DCNN models using three benchmark data sets for the fine-grained image classification task demonstrate the effectiveness of our method.

8. REFERENCES

[1] J. Krause, M. Stark, J. Deng, and L. Fei-Fei, “3D object representations for fine-grained categorization,” in Proc. IEEE Int. Conf. Comput. Vis. Workshops, Jun. 2013, pp. 554–561.

[2] Y.-L. Lin, V. I. Morariu, W. Hsu, and L. S. Davis, “Jointly optimizing 3D model fitting and fine-grained classification,” in Proc. Eur. Conf. Comput. Vis., 2014, pp. 466–480.

[3] L. Yang, P. Luo, C. Change Loy, and X. Tang, “A large-scale car dataset for fine-grained categorization and verification,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2015, pp. 3973–3981.

[4] X. Zhang, F. Zhou, Y. Lin, and S. Zhang, “Embedding label structures for fine-grained feature representation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 1114–1123.

[5] F. Zhou and Y. Lin, “Fine-grained image classification by exploring bipartite-graph labels,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 1124–1133.

[6] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie, “The caltech-UCSD birds-200-2011 dataset,” Dept. Comput. Neural Syst., California Inst. Technol., Pasadena, CA, USA, Tech. Rep. CNS-TR-2011-001, 2011.

[7] T. Berg, J. Liu, S. Woo Lee, M. L. Alexander, D. W. Jacobs, and P. N. Belhumeur, “Birdsnap: Large-scale fine-grained visual categorization of birds,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 2011–2018.

[8] S. Branson, G. Van Horn, C. Wah, P. Perona, and S. Belongie, “The ignorant led by the blind: A hybrid human–machine vision system for fine-grained categorization,” Int. J. Comput. Vis., vol. 108, nos. 1–2, pp. 3–29, 2014.

[9] S. Li, K. Li, and Y. Fu, “Self-taught low-rank coding for visual learning,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 3, pp. 645–656, Mar. 2018.

[10] S. Maji, E. Rahtu, J. Kannala, M. Blaschko, and A. Vedaldi. (2013). “Fine-grained visual classification of aircraft.” [Online]. Available: https://arxiv.org/abs/1306.5151

[11] T.-Y. Lin, A. RoyChowdhury, and S. Maji, “Bilinear CNN models for fine-grained visual recognition,” in Proc. IEEE Int. Conf. Comput. Vis., Dec. 2015, pp. 1449–1457.

[12] X. Zhao et al., “Scalable linear visual feature learning via online parallel nonnegative matrix factorization,” IEEE Trans. Neural Netw. Learn. Syst., vol. 27, no. 12, pp. 2628–2642, Dec. 2016.

[13] L. Niu, W. Li, D. Xu, and J. Cai, “An exemplar-based multi-view domain generalization framework for visual recognition,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 2, pp. 259–272, Feb. 2018.

[14] T. Rumbell, S. L. Denham, and T. Wennekers, "A spiking self-organizing map combining STDP, oscillations, and continuous learning," IEEE Trans. Neural Netw. Learn. Syst., vol. 25, no. 5, pp. 894–907, May 2014.

[15] C. Zhang, C. Liang, L. Li, J. Liu, Q. Huang, and Q. Tian, "Fine-grained image classification via low-rank sparse coding with general and classspecific codebooks," IEEE Trans. Neural Netw. Learn. Syst., vol. 28, no. 7, pp. 1550–1559, Jul. 2017.

[16] H. Zhang, L. Cao, and S. Gao, "A locality correlation preserving support vector machine," Pattern Recognit., vol. 47, no. 9, pp. 3168–3178, Sep. 2014.