Analysis of Learning Techniques: Bird Species Classification

Pranav Kumar Rai
Student
Galgotias College of Engineering and Technology
Knowledge Park-1, Greater Noida

Sourvi Chaturvedi
Student
Galgotias College of Engineering and Technology
Knowledge Park-1, Greater Noida

Sandhya Katiyar, PhD
Associate Professor
Galgotias College of Engineering and Technology
Knowledge Park-1, Greater Noida

ABSTRACT

Identification of bird species is a difficult problem that pushes the limits of the visual abilities for both humans and computers. Although different bird species share the same basic set of parts, different bird species can vary dramatically in shape and appearance[5]. Sometimes professional bird watchers disagree on the species given an image of a bird. Intra-class variance is high due to variation in lighting and background and extreme variation in pose (e.g., flying birds, swimming birds, and perched birds that are partially occluded by branches).

In this paper, a simple image recognition classifier has been created. This image recognition tool classifies various species of birds. An application CNN has been used in order to extract features from the input image. Convolution preserves the spatial relationship between pixels by learning image features using small squares of input data. Then application of some of the supervised and unsupervised algorithms is used to check and compare their accuracies against each other.

The aim is to find which learning algorithm's accuracy is best in predicting a particular specie of bird given any image of it.

General Terms

Bird Species Classification , Image Classification , Pattern Recognition ,CNN , Learning algorithms.

Keywords

CNN, Image classifier, accuracy, bird species.

1. INTRODUCTION

Identification of bird species is a challenging task often resulting in ambiguous labels. Convolutional neural networks (CNN) is a special architecture of artificial neural networks, proposed by Yann LeCun in 1988. CNN uses some features of the visual cortex. One of the most popular uses of this architecture is image classification[1]. Two types of learning algorithms have been applied to train the data and test their accuracy against each other in predicting the species of the birds. It was observed that libraries like Scikit Learn allowed to tweak different aspects of an algorithm, but may be not to the extent of   implementation of the algorithm and faced an inherent trade-off between tweaking ability and the number of algorithms that could be implemented and tested in the time frame of the research. The research was about trying out numerous algorithms using the Scikit Learn library[2] in Python:

1. Naive Bayes

2. Support Vector Machines

3. K-nearest Neighbors

4. Linear Discriminant Analysis (LDA)

5. Decision Trees

6. Random Forests

7. One versus Rest classifiers with Logistic Regression[4]

1.1 Supervised Learning algorithms
This paper describes 6 types of supervised algorithms to check the accuracy of predicting the species of birds.

1. Naive Bayes
Naïve bayes algorithm is used for classifier construction. These classifiers contain models that give class labels to problem statement and these problem statement are represented as vectors with some feature values. All naïve bayes classifier take the value of a particular feature as independent value. Below is an equation depecting a Bayesian probability distribution.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

P(c|x) $\longrightarrow$ Posterior probability

P(x|c) $\longrightarrow$ Likelihood

P(x) $\longrightarrow$ Predictor prior probability

P(c) $\longrightarrow$ Class prior probability

2. Support Vector Machines
It is also known as support vector networks. It is a method used in supervised problems like classifications. The original SVM algorithm was developed by Vladimir N. Vapnik. It follows two ideas, firstly linear seperable classes and secondly separting hyperplanes with maximum margins.

Considering a trining set : (x,y), i = 1,2…N; y $\mathcal{E}$ {+1,-1}

Hyperpane: wx + b = 0
w = $w_1, w_2…wd$ *and* $x= x1, x2….xd$

3. K-nearest Neighbors
It is the extension of nearest neighbor rule. It classifies a feature value "x" by giving it, the label that have been frequently represented among the k-nearest sample. If k is fixed and the given number of samples is allowed to approach to infinity then all the k-nearest neighbors will converge to x.

Euclidean distance function $\left(\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}\right)$

Manhattan $\sum_{i=1}^{k}|x_i - y_i|$

4. *Decision Trees*

A decision tree is a pre-arrangement of tests that provides a predefined classification at every step during an analysis. Moreover, decision trees classify these tests by storing them down the tree to a root node following some leaf node, that provide the classification of the test. The decision tree includes a root node, and several leaf nodes which are a set of possible answers of the condition of root node.
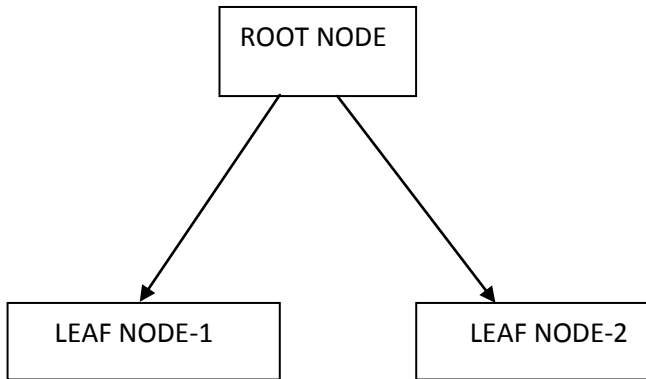
5. *Random Forests*

Random forest also known as semi-ensemble learning method. The method of random forest was first proposed by HO in 1995[13]. This learning technique is another method of classification, regression. It collect results of different decision trees and calculates their mean prediction as a final result. This method corrects the overfitting problem of decision trees to their training set.

$$RFfi_j = \frac{\sum_j normfi_{ij}}{\sum_{j \varepsilon allfeatures, k \varepsilon alltrees} normfi_{jk}}$$

6. *One versus Rest classifiers with Logistic Regression*

One versus rest is known as one-versus-all or one-against-all. In this classify two different classes of objects from given set of data(in which the classes of data is classified together). Similarly all the other classes tested against each other individually and their results are considered positive and negative according to the class under consideration. It support decision based classification .
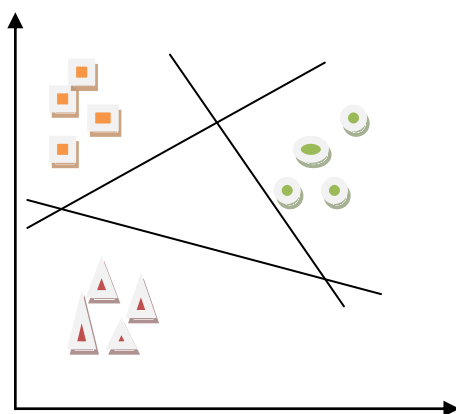

Figure:1 one vs rest logistic regression

1.2 *Unsupervised Algorithms*

LDA an unsupervised algorithm have been applied to check the accuracy of predicting the species of birds. Recent publications suggest that unsupervised pre-training of deep, hierarchical neural networks improves supervised pattern classification.[9,10]

i. *Linear Discriminant Analysis (LDA)*

Linear discriminant analysis (LDA) is used in statistics, pattern recognition and machine learning, this method separates the two or more classes of object by making use of linear combination of features. It also used in dimensionality reduction before applying classification later.

$$\beta = C^{-1}(\mu_1 - \mu_2)$$

$$C = \frac{1}{n_1 + n_2}(n_1 C_1 + n_2 C_2)$$

β : linear model coefficient , μ_1, μ_2 : mean vector

C_1, C_2 : covariance matrix , C : pooled covariance matrix

LDA is closely related to analysis of variance (ANOVA) and regression analysis, which also attempt to express one dependent variable as a linear combination of other features or measurements.[14][15]

2. LITERATURE SURVEY

Aditya Bhandari, Ameya Joshi, Rohit Patki in their paper have described their implementation using KNN and Naïve Bayes in MATLAB. They have used physical features of birds for their classification.[4]

In an another paper *Cheng Pang, Hongxun Yao, Xiaoshuai Sun* have done their implementation using discriminative features and pattern recognition technique.[5]

John Martinsson in his paper 'Bird species identification using convolutional neural networks' aimed to improve upon the state of the art of the bird species classifiers. He used deep residual neural networks to classify bird species.

Rosniza Roslan, Nursuriati Jamil in their paper have done classification and Identification of bird species using visual image. They have applied the knowledge of feature species type. They have done classification based on nine colour based features of mean, 'standard deviation and skewness of each plane of red, green and blue(RGB) from the bird images.

Diego Rafael Lucio, Yandre Maldonado, Gomes da Costa in their paper described a system for automatic bird species classification based on feature taken from the textural content of spectrogram images. The experiments were performed on their exclusive data set containing 46 different species and the best accuracy rate observed was about 77.65%.

Akash Gupta, Sourya Dipta Das in 'Bird species classification using transfer learning with multistage training' have researched about image classification using computer vision. A pre-trained mask- iRCNN and a ensemble model consisting of inception of nets have been used to get the both localization and species of bird from the images.

Andreia Marini, Jacques Facon , Alessandro L Koerich in their research paper have followed a anoivel approach for bird species classification. The classification was based on colour feature extracted from un constrained images. This experiment yielded results on the CUB-200 dataset with an accuracy of about 75% of correct segmentation rate.

Juha Niemi, Juha T. Tanttu in ' Deep learning case study for automatic bird identification' used visual camera images as external data apart from data from radar for automatic bird detection. A convolutional neural network is used with trained

data and a deep learning algorithm have been applied for classification.

R.Yoshihashi, R.Kawakami, M. Lida, T. Naemura in 'bird detection and species classification with time-lapse images around a wind farm:data construction and evaluation' have talked about automatic bird monitoring that can be of aid resolving collision of birds.

Grant Van Horn, Steve Branson,Ryan Farrell , Scott Haber in 'Building a bird recognition app and large scale dataset' have used tools and methodologies to collect high quality, large scale computer vision datasets.

Chao Zhang , Takuya Akashi in 'Cross-domain deep feature combination for bird species classification with audio-visual data' have proposed CNN based multi-model learning models in three types of fusion strategies i.e. early, middle and late, to settle the issue of combining data cross domains

3. PROPOSED FRAMEWORK

3.1 Image Classification

This paper is an introduction to machine learning with scikit-learn, a popular and well-documented Python framework. This research uses Python 3 to build an image recognition classifier which accurately determines the bird species displayed in images . Convolutional Neural Networks (CNNs) is the most popular neural network model being used for image classification problem. The big idea behind CNNs is that a local understanding of an image is good enough. The practical benefit is that having fewer parameters greatly improves the time it takes to learn as well as reduces the amount of data required to train the model. Instead of a fully connected network of weights from each pixel, a CNN has just enough weights to look at a small patch of the image.

3.2 Dataset

Caltech and UCSD have gathered data to produce the "Caltech-UCSD Birds-200-2011 (CUB-200-2011)" dataset[3]. The dataset contains 11,788 images of 200 bird species. The list of species names was obtained using an online field guide. Images were harvested using Flickr image search and then filtered by showing each image to multiple users of Mechanical Turk. Caltech have verified the effectiveness of the method on the subsets of ILSVRC2012 dataset and CUB200 2011 dataset.[8]There are totally 200 species of birds in the dataset, it was observed that the samples from a certain species of bird is sparse among the overall dataset. It can be inferred that one can represent a pairwise parts using a few bases acquired from the same pairwise parts of other samples. Since sparse representation is shown to be effective local image descriptors which could capture subtle patterns and lead to less quantization errors.

Figure: 2 Caltech-UCSD Birds 200 dataset. (Caltech.edu)

3.3 Tools and Technologies

A. **ANACONDA-** Anaconda is a free and open source distribution of the Python and R programming languages for data science and machine learning related applications , that aims to simplify package management and deployment. [6].

B. **TensorFlow-** TensorFlow is an open-source software library for dataflow programming across a range of tasks.

C. **KERAS -** Keras is an open source neural network library written in Python. Activate Tensorflow and install keras using 'pip install keras'

D. **CNN -** Convolution Neural network , a class of deep, feed-forward artificial neural networks, most commonly applied to analyzing visual imagery

3.4 Building the CNN

This is most important step for our network. It consists of three parts –

1. Convolution
2. Polling
3. Flattening

The primary purpose of CNN is to extract features from the input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data.
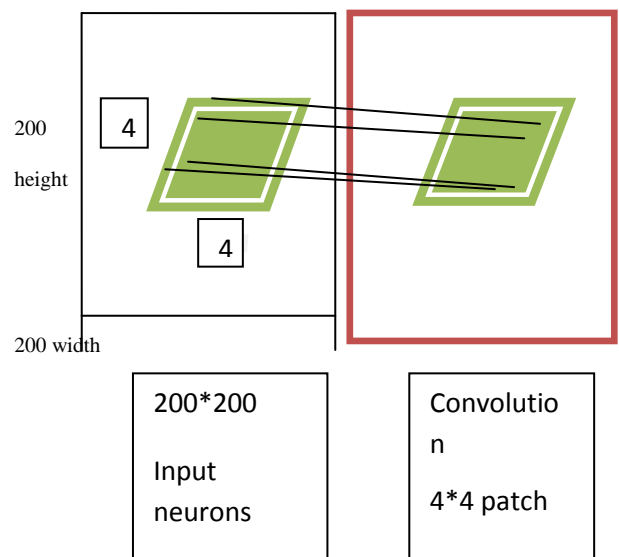
Consider a 200 x 200 image. CNN can efficiently scan it chunk by chunk—say, a 4×4 window. The 4×4 window slides along the image (usually left to right, and top to bottom), as shown below. How "quickly" it slides is called its stride length. For example, a stride length of 2 means the 4×4 sliding window moves by 2 pixels at a time until it spans the entire image. The neural network, which has 60 million parameters and 500,000 neurons, consists of five convolutional layers, some of which are followed by maxpooling layers.A convolution is a weighted sum of the pixel values of the image, as the window slides across the whole image[7]. Turns out, this convolution process throughout an image with a weight matrix produces another image (of the same size, depending on the convention). Convolving is the process of applying a convolution. the pixel being interpolated

is first classified in the context of a window of neighboring pixels; and then the corresponding high-resolution pixels are obtained by filtering with coefficients that depend upon the classification.[11]. A typical CNN has multiple convolution layers. Each convolutional layer typically generates many alternate convolutions, so the weight matrix is a tensor of 4 × 4 × b, where b is the number of convolutions. The segmentation procedure of region merging is implemented as a hierarchical clustering algorithm.[12].

3.5 Training and Testing

The dataset was trained and tested. But before it was done, a need to split the total collection of images into two sets – one for training and one for testing and also added a third set for development/validation. Keeping the testing set completely separate from the training set is important, because it needed to be sure that the model will perform well in the real world. Once trained, it will have seen many example images of species of birds. It was needed to be the matter of surity that when presented with new images of birds it hasn't seen before, that it has actually learnt something from the training and can generalise that knowledge – not just remember the exact images it has already seen. Usually, between 70-90% of the data for training was used, though this varies depending on the amount of data collected, and the type of model trained.
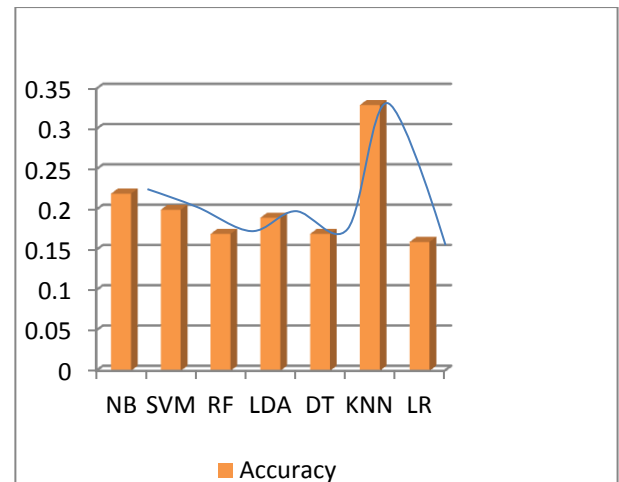
3.6 Statistical Analysis

Table show the results that were observed on feeding dataset to following algorithms as mentioned in the above section. Here it is clearly visible that KNN posses the best accuracy among all i.e. 0.33. As naïve bayes is second best algorithm for used data set, it shows its accuracy about 0.22. SVM can be seen performing with accuracy 0.2 which is nearly equal to Naïve bayes. Algorithms like LDA, Random forest and Decision tree shows nearly equal predictions with accuracies 0.19, 0.17, 0.17 respectively. The One vs Rest Logistic Regression algorithm perform with the least accuracy of 0.16. If the training data is to be increased then these same algorithms may yield better results.

Table 1 : Algorithms and their calculated accuracies.

Algorithms	Training Accuracy
Naïve Bayes	0.22
SVM	0.2
Random Forest	0.17
LDA	0.19
Decision Tree	0.17
KNN	0.33
One vs rest – Logistic	0.16

4. RESULT

Following is the graphical representation of accuracies of all the algorithms that we have used for the classification of the species of birds.

Y-axis: accuracy and X-axis: applied algorithms

The above is the graphical representation of the statistical analysis of all the applied learning algorithms for the dataset used in this paper for applying bird classification. The bars in the graph shows the variation in the accuracies of each of the algorithms when tested under same dataset values. The curved line over the bars show the points of variation for each algorithm on the graph. It is clear while reading the graph that KNN algorithm outperformed all other algorithms used in the paper while Logistic Regression(one vs rest) is least efficient in classification for the given dataset and shows the lowest accuracy value.

5. CONCLUSION

Learning Algorithms were trained and tested on the complete data set to start with. Later randomly into training data and test data was done so that dataset had samples from each class. 70% of the data was used as training data and 30% was used as test data. It was found that KNN predicts more accurately the images for the given dataset

6. FUTURE WORK

1. Implementation of Neural Networks was performed and when ran it on the machine for just 5 hidden neurons, it went out of memory and could not complete. So, one can try to run Neural Networks on high performance computing machines.

2. Computer vision algorithms can be used for automatic feature extraction.

3. One can develop an Android/iOS application that identifies a bird in real time on clicking its photo.

7. ABBREVIATIONS

NB- Naïve Bayes

SVM- Support Vector Machine

RF – Random Forest

LDA- Linear Discriminant Analysis

DT- Decision Tree

KNN– K- Nearest Neighbor

LR – One vs Rest Logistic Regression

8. REFERENCES

[1] Steve Branson et al. "Bird Species Categorization Using Pose Normalized Deep Convolutional Nets". In: CoRR abs/1406.2952 (2014). URL: http://arxiv.org/abs/1406.2952.

[2] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: Journal of Machine Learning Research 12 (2011), pp. 2825–2830.

[3] C. Wah et al. "The Caltech-UCSD Birds-200-2011 Dataset. Tech. rep. CNS-TR-2011-001". California Institute of Technology, 2011.

[4] "Bird Species Identification from an Image"Aditya Bhandari, Ameya Joshi, Rohit Patki1Department of Computer Science, Stanford University2Department of Electrical Engineering, Stanford University3Institute of Computational Mathematics and Engineering, Stanford University

[5] Discriminative Features for Bird Species ClassificationCheng Pang Hongxun Yao Xiaoshuai SunSchool of Computer Science & Technology, Harbin Institute of Technology, China{pangcheng3, h.yao, xiaoshuaisun}@hit.edu.cn

[6] *Chao Zhang , Takuya Akashi ,* "Cross-domain deep feature combination for bird species classification with audio-visual data", August 2017

[7] "ImageNet Classification with Deep Convolutional Neural Networks"Part of: Advances in Neural Information Processing Systems 25 (NIPS 2012) Alex Krizhevsky , Ilya Sutskever, Geoffrey E. Hinton

[8] The Application of Two-level Attention Models in Deep Convolutional Neural Network for Fine-grained Image Classification Tianjun Xiao1 Yichong Xu2 Kuiyuan Yang2 Jiaxing Zhang2 Yuxin Peng1∗ Zheng Zhang3 1 Institute of Computer Science and Technology, Peking University

[9] A. C. P.-A. M. P. V. Dumitru Erhan, Yoshua Bengio and S. Bengio. Why does unsupervised pretraining help deep learning? Journal of Machine Learning Research, 11:625–660, 2010.

[10] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle. Greedy layer-wise training of deep networks. In Neural Information Processing Systems, 2007

[11] Optimal image scaling using pixel classificationC.B. Atkins ; C.A. Bouman ; J.P. Allebach

[12] Unsupervised multistage image classification using hierarchical clustering with a bayesian similarity measureSanghoon Lee ; M.M. Crawford

[13] Ho, Tin Kam (1995). *Random Decision Forests* (PDF). Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14– 16 August 1995. pp. 278–282. Archived from the original(PDF) on 17 April 2016. Retrieved 5 June 2016.

[14] Garson, G. D. (2008). Discriminant function analysis. https://web.archive.org/web/20080312065328/h ttp://www2.chass.ncsu.edu/garson/pA765/discrim.htm.

[15] Jump up to: Hardle, W., Simar, L. (2007). *Applied Multivariate Statistical Analysis*. Springer Berlin Heidelberg. pp. 289-303.