IJCSI International Journal of Computer Science Issues, Volume 17, Issue 1, January 2020
ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784
www.IJCSI.org https://doi.org/10.5281/zenodo.3987114

Classification of Birds Using KNN and SVM Classifier

Naveena M[1], G Hemantha Kumar[2]

**[1] Scientific officer, Manasagangothri
University of Mysore, Mysore**

**[2] Professor and Vice Chancellor
University of Mysore, Mysore**

Abstract:
This paper aims to develop a bird's classification system based on classifiers fusion to easily identify the birds. It is based on image processing, which can control the classification, qualification and segmentation of images and hence recognize the birds. Usually from the captured images multiple shape features like area, centroid, angle at centroid, maximum angle and minimum angle can be extracted and analyzed to classify and recognize the birds. And then the extracted features are classified using KNN (K-Nearest Neighbor) Classifier and SVM (Support vector machine) classifiers.

1. INTRODUCTION

Bird species identification from images is an important and challenging problem with many applications in the real world. In order to evaluate the quality of our living environment it is important to obtain reliable information about the population of wild animals. Birds are numerous and sensitive to environmental changes; also, and are easier to monitor than other species. Therefore, the use of automated methods for bird species identification is an effective way to evaluate the quantity and diversity of the birds. Bird identification is a well-known problem to ornithologists, and is considered as a scientific task since antiquity. Ornithologists study birds; their existence in nature, their biology, their songs, their distribution, and their ecological impact. Bird classification is usually done by ornithology experts based on an animal classification system proposed by Linnaeus: Kingdom, Phylum, Class, Order, Family, and Species. Birds are typically categorized by their shape or color and physical characteristics. Bird species classification is a challenging problem both to humans and to computational algorithms because this problem is based on bird's images. Compared with sound identification, visual features are not well studied for bird classification. The classification problem can be stated as given a bird image, classify its species among a fixed but large number of possibilities. The challenge of such a classification task is due to the variation in the background and illumination since most of the images are gathered on birds' natural habitat and in the birds pose since its not possible to control rotation, scale, and angle of view while acquiring images. Visual properties (e.g., shape, color, marks, etc.) are important keys for bird recognition. Some researchers utilize these features to automatically identify birds. They compared several image processing techniques for the classification of bird's species. Therefore, it is particularly important to classify the birds. With the development of image identification technology, using modern technology to identify birds has become an effective research method.

Bird's play an important roles in the ecosystem, e.g. providing insect control and seed dispersion and pollination. As a result, it is vital to improve knowledge about a bird's identify, species, evolution, classification, and geographic distribution. This project introduces a new approach of bird's classification which is a part of image processing. Here classification is done using KNN and SVM classifiers by using the some bird's images as a dataset.

A machine vision system to identify and classify birds, a substantial work dealing with the use of features set for classification of different types of birds will be proposed. For bird's classification methodology images are resized and converted into grayscale image, edge is detected from the grayscale images and features are extracted. Then, they are fed to various classifiers to classify them accordingly. The various classifiers included are Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) algorithm. The purpose of the K- Nearest Neighbors (KNN) algorithm is to use a database in which the data points are separated into several separate classes to predict the classification of a new sample point. The SVM classifier classifies the data by mapping the vector from low-dimensional space to high-dimensional space using kernel function.

At last fuse the results obtained from both classifiers by using logical or operator and get classified label.

2. PROPOSED METHODOLOGY

In this project we proposed a new methodology for birds classification based on fusion with KNN(K-nearest neighbor) and SVM (Support Vector machine) classification methods. We take bird images as training images to train the system first, then we resize an original image, then we do the edge detection methodology for a binary image using 'sobel' method by doing this we can easily extract the features without any noise, and store those features in the database. Then take one testing image and do the same preprocessing steps for the testing image, net step is to compare features (values) obtained from testing bird image with the database values using KNN and SVM classification methods and finally get the classified label of bird from both classification methods and compare the accuracies of both methodologies, which shows which bird belongs to which group in the training sets.

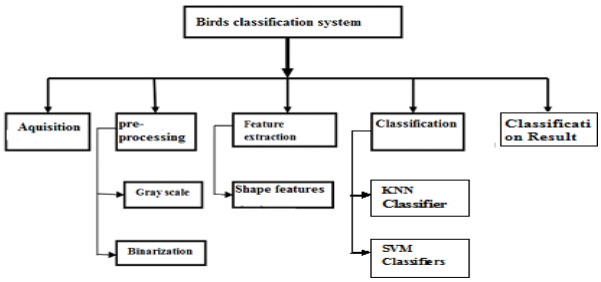

Fig1: Proposed Architecture

3. FEATURE EXTRACTION

A feature is defined as an "interesting" part of an image, and is used as a starting point in main primitives for subsequent algorithms. The overall algorithm will often only be as good as its feature detector. Consequently, the desirable property for a feature detector is its repeatability: whether or not the same feature will be detected in two or more different images of the same scene. The most important types of features which can be considered when trying to identify the signs are spatial, temporal and textural. The feature extraction stage is built and designed to process real images .The algorithms used in these systems are commonly divided into three tasks: extraction, selection and classification.

In image processing feature extraction is one of the important steps; it is a type of dimensionality reduction. When the input data is too large to be processed and suspected to be redundant then the data is transformed into a reduced set of feature representations. The process of transforming the input data into a set of feature is called feature extraction, feature contain information relative to color, shape, texture or context. In this project we are going to use the shape feature like area, centroid, maximum angle minimum angle etc.

3.1 Shape feature extraction

The shape of an image is refers to its physical structural and profile. Shape feature is used to finding and matching shapes of an image, and recognizing or making measurements of shapes moment, perimeter, area and orientation are some of the characteristics used for shape feature extraction techniques. But in this our project finding the area, centroid, angle at centroid, maximum angle and minimum angle of the bird image to extract a shape feature.

Using shape feature we have to classify the various birds species. In shape feature we are calculate the area, centroid, angle at centroid, maximum angle and minimum angle of the bird image using MATLAB code.

3.2 Area (bwarea):

Area of objects in binary image Estimate the area of the objects in binary image BW. And overall result is a scalar whose value corresponds roughly to the total number of on pixels in the image but might not be exactly the same because different patterns of pixels are weighted differently.

3.3 Centroid:

Centroid of the objects in binary image Vector that specifies the center of the region. Note that the first element of Centroid is the horizontal coordinate (or x-coordinate) of the center of mass, and the second element is the vertical coordinate (or y-coordinate). All other elements of Centroid are in order of dimension.

3.4 Angle at Centroid

It is used to identify the angle which is at the centroid, it is one of the useful future used in this project for the classification of the bird's species .Here we find the value of the angle which is located at the centroid of the image which is converted to the binary image.

3.5 Maximum and minimum angle

It refers to the angles from the centroid of the image, here we can find two angle like maximum angle and minimum angle .Maximum angle refers to the angle

which has maximum values at the centroid. Minimum angle refers to the angle which has minimum value at the centroid. These features are more helpful to identify the shape of the birds.

4. CLASSIFICATION

Classification is a process in which individual items are grouped based on the similarity between the item and description of the group. Classification includes a broad range of decision-theoretic approaches to the identification of images. Classification algorithm typically employs two phases of processing training and testing. Classification is the process by which we attribute a class label to set of measurements essentially. This is the heart of the pattern recognition. There must be many approaches. These include statistical and structural approaches, among the various method of supervised statistical pattern recognition. The Nearest Neighbor rule achieves consistently high performance. Without apriori assumptions about the distributions from which the training examples are drawn. It involves a training set of both positive and negative class.

EXPERIMENTAL RESULTS AND OBSERVATIONS

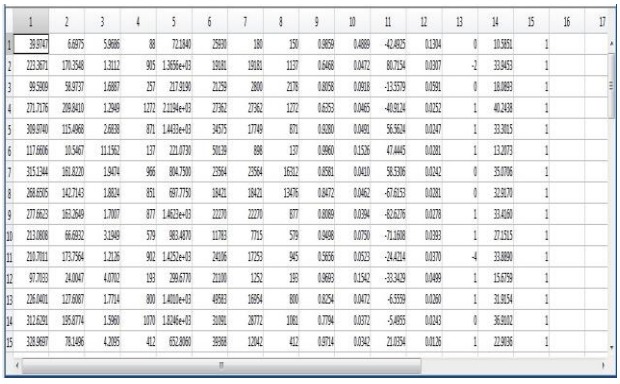

Fig1: Extracted Feature values

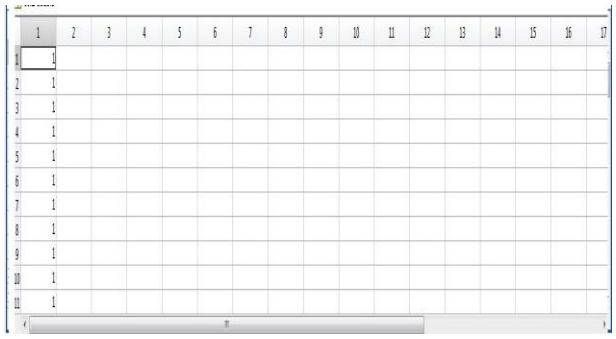

Fig2: Labeling the images

RESULT AND ANALYSIS

4.1 KNN Classifier

KNN classification is a method which classifies the data provided by considering the closest training values of objects by comparing with testing samples. In this table we have considered two values those are ground truth values and no. of correctly classified samples. So that we can easily get the accuracy for each class. Finally we got a 52% accuracy after comparing the results of ground values and no of correctly classified samples.

Train%-Test%	Accuracy	Precision	Recall	f-measure
80-20	53	100	75	85.7143
70-30	55	40	100	57.14
60-40	41.67	25	33.333	28.5714
50-50	46.66	50	40	44.444
40-60	44.44	37.50	66.66	48.00

4.2 SVM Classifiers

The supporting vector machine (SVM) is a supervised learning method that generates input-output mapping function from a set of labeled training data. The mapping function can be a classification function, i.e., the category of input data. For classification we are considering two values those are ground values and no of correctly classified samples. After that we will match the results from two values and find the percentage of accuracy. Here also we have taken 5 classes with 20 samples for each class as we have taken in KNN. After analyzing the results of both ground values and no of correctly classified samples.

Train%-Test%	Accuracy	Precision	Recall	f-measure
80-20	72.2222	50	66.6666	57.1429
70-30	66.6667	56	33.3333	40
60-40	70.8333	100	100	100
50-50	60	33.3333	25	28.5714
40-60	55	41.6667	55.555	47.6190

5. Comparative analysis:

By analyzing the methodologies of KNN and SVM we can compare the accuracy values for both classifiers. Here we proposed a methodology to classify the bird images by using KNN and to obtain the accuracy rate, and do the same preprocessing procedure for SVM classification for the same features. The results and observations show that SVMs are a more reliable more of classifiers. However, KNN is less computationally intensive than SVM. Since, KNN is easy to implement, the classification of Multi-class data should be done with KNN. The algorithm that guarantees reliable detection in unpredictable situations depends upon the data. If the data points are heterogeneously distributed, both should work well. If data is homogenous to look at, one might be able to classify better by putting in a kernel into the SVM. For most practical problems, KNN is a bad choice because it scales badly - if there are a million labeled examples, it would take a long time (linear to the number of examples) to find K nearest neighbors.

6. CONCLUSION

In this work, methodology of classification of birds based on the classifiers fusion we used KNN (K-nearest neighbors) and SVM (Support Vector Machine) is used. Apart from this, the bird's recognition system which tells which sample bird belong to which class in database. It is concluded that, it is very challenging task to build up an automatic bird classification system that works effectively in all situation weather the bad image condition. Here we have used 5 different classes of birds images and extracted the shape feature like area, centroid, maximum angle and minimum angle all these features are used to train the system to identify the bird's images .For the classification we have used classifiers fusion by using two classifiers i.e. KNN and SVM classifiers and compare the results of those classifiers to find the best results Using the logical or operator we find the best of the classification result among the those classifier. Most of the existing methodology and algorithms are unable to provide a solution for those problems of variation.

FUTURE WORK

The goal, robust bird's classification is still far. The bird's classification based on classifiers fusion can be implemented by using more than two classifiers to identify the birds but while classifying the birds images we can do for the images which are have complex background because image segmentation itself is a challenging task in birds classification. Here we have did classification for the images which have simple background .Classification of bird's by combining two or more than that has been a challenging task due to the number of factors like image condition and illumination. Most of the existing methodology and algorithms are unable to provide a solution for those problems of variations. A considerable attention is required to develop efficient methods that can work with and overcome to such problem and to get a different methodology and to get more accuracy other classification methods can be applied.

REFERENCES

[1] U. D. Nadimpalli, R. R. Price, S. G. Hall, and P. Bomma, "A comparison of image processing techniques for bird recognition," Biotechnology Progress, vol. 22, no. 1, pp. 9–13, 2006.
[2] E. P. Kasten, P. K. McKinley, and S. H. Gage, "Ensemble extraction for classification and detection of bird species," Ecological Informatics, vol. 5, no. 3, pp. 153 – 166, 2010. [Online]. Available:http://www.sciencedirect.com/science/article/pii/S1574954110000361
[3] T. S. Brandes, "Automated sound recording and analysis techniques for bird surveys and conservation," Bird Conservation International, vol. 18, pp. 163–173, 2008.
[4] S. Fagerlund, "Bird species recognition using support vector machines," EUSASIP J. Adv. Signal Process., vol. 2007, pp. 1–8, 2007.
[5] H. Lin, H. Lin, and W. Chen, "Study on recognition of bird species in minjiang river estuary wetland," Procedia Environmental Sciences, vol. 10, Part C, pp. 2478 – 2483, 2011.
[6] M. Das and R. Manmatha, "Automatic segmentation and indexing in a database of bird images," in Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on, vol. 2, 2001, pp. 351 –358 vol.2.
[7] H.-C. Wang, Y.-S. Chen, and M.-Y. Wu, "A user-augmented object query system using color and shape features for taiwan wild birds photos," in Machine Learning and Cybernetics (ICMLC), 2010 International Conference on, vol. 5, july 2010, pp. 2516 –2520.
[8] V. Vapnik, Statistical Learning Theory, New York: John Wiley & Sons,1998, 732 p.
[9] O. Chapelle, V. Vapnik, O. Bousquet, S. Mukherjee, Choosing Multiple Parameters for

Support Vector Machine, Machine Learning, 46, 2002,pp. 131-159.
[10] Lean Yu, Shouyang Wang, Kin Keung Lai, Ligang Zhou, Bio-Inspired Credit Risk Analysis. Computational Intelligence with Support Vector Machines, Springer-Verlag Berlin Heidelberg, 2008, 244 p.
[11] L. Demidova, E. Nikulchev, Yu. Sokolova, Use of Fuzzy Clustering Algorithms Ensemble for SVM classifier Development, IREMOS, vol. 8, no. 4, pp. 446-457, August 2015.
[12] L. Demidova, Yu. Sokolova, Training Set Forming For SVM Algorithm With Use Of The Fuzzy Clustering Algorithms Ensemble On Base Of Cluster Tags Vectors Similarity Matrices. International Conference "Stability and Control Processes" in Memory of V.I. Zubov, pp. 619-622, October 2015.
[13] L. Demidova, Yu. Sokolova, Modification Of Particle Swarm Algorithm For The Problem Of The SVM Classifier Development. International Conference "Stability and Control Processes" in Memory of V.I. Zubov, pp. 623-627, October 2015.
[14] L.A. Demidova, E.V. Nikulchev, Yu.S. Sokolova The SVM Classifier Based On The Modified Particle Swarm Optimization, International Journal of Advanced Computer Science and Applications (IJACSA), vol. 7, no. 2, pp. 16-24, 2016.