MACHINE LEARNING MINI - PROJECT

FLOWER SPECIES CLASSIFICATION

Ankith Reddy Adudodla - 1901013
Indian Institute of Information Technology Guwahati

Outline

- Problem Statement
- Abstraction
- Introduction
- Literature Survey
- Results Analysis

Problem Statement

Classifying the flowers into one of the 17 categories or species based on colour, shape and texture using various Machine Learning Algorithms and comparing their performances.

Abstraction

- Recently, flower classification is attracting more and more attention, which is often used in floriculture, flower searching or patent analysis, etc. Identifying flower categories manually is tedious and time consuming. Even experienced botanists need reference books when identifying the categories of the flowers. Therefore, developing automatic flower classification systems is necessary and urgent, which can identify the flower categories fast and accurately.
- To this end, in this mini-project, we made a comparison of different classification algorithms such as Logistic Regression, PLA, Single Layer Perceptron and MultiLayer Perceptron, which is evaluated on a very challenging database of flower images.
- Results are presented on a Oxford dataset of 1360 images consisting of 17 flower species by considering various combinations of colour, shape, scale and viewpoint.
- In this project we observed that Logistic Regression (80 %) has a better accuracy compared to another ML algorithms.

Introduction

- It’s a challenging task to classify the flowers because of large intra-class differences and relatively small inter-class differences and also some flowers are need to be identified by only dissecting them.
- Foreign scholars, Nilsback M.E and Zisserman A once proposed a visual word package model. After the flower image is segmented, the texture, color and other features of the flower image are extracted, and these extracted features are imported into the classifier to complete the classification of the flower image.
- To improve the visual experimental packet model, they created the Oxford 17 flowers’ data set also, so in this data set , there are 17 types of flower with 80 images in each.
- In this project the results are presented based on the extracted features HSV, HOG, SIFTINT, SIFTBDY.

- The oxford flowers dataset consists of 17 different flower species namely :

1. Buttercup
2. Sunflower
3. Snowdrop
4. Coltsfoot
5. Daffodil
6. Daisy
7. Dandelion
8. Fritillary
9. Iris
10. Pansy
11. Windflower
12. LilyValley
13. Bluebell
14. Crocus
15. Tigerlily
16. Tulip
17. Cowslip

- Each of the five features provided in the dataset contains a 2D matrix of 1360 * 1360 shape. This matrix contains 1360 patterns (17 species * 80 flowers in each in the order) and 1360 extracted sub-features or pixels.

- The features that are extracted and provided are :

  1. **Color :** It is described by taken the **HSV** values of the pixels.The HSV values for each pixel in an image are clustered using k-means.

  2. **SIFT (Scale Invariant Feature Transform):** The SIFT features (SIFTINT - **SIFINT** Internal and **SIFTBDY** - SIFT Boundary) describe both the texture and the local shape of the flower (e.g. fine petal structures (such as a sunflower) vs spikes (such as a globe thistle).

  3. **Histogram of Gradients: HOG** features are similar to **SIFT** features, except that they use an overlapping local contrast normalization between cells in a grid.

Literature Survey

Automated flower classification over a large number of classes Year 2008 by Nilsback and Zisserman

- In this paper they computed four different features for the flowers, each describing different aspects, namely the local shape/texture, the shape of the boundary, the overall spatial distribution of petals, and the colour. [HSV, SIFTINT, SIFTBDY, HOG]
- They combined the features using a multiple kernel framework with a SVM classifier. A weighted linear combination of kernels is used, one kernel corresponding to each feature.
- The final kernel has the following form for two data points i and j: $K(i,j) = \sum_{f \epsilon F} \beta_f \exp\left(-\mu_f \chi_f^2(x_f(i), x_f(j))\right)$
- They have experimented and observed that by combining all the features extracted, they were able to achieve a better performance compared to single feature.
- They are able to achieve an accuracy of **88.33 ± 0.3%**
- The beside table shows the experimental results made by them In 2006 by **nearest neighbour classifier.** They observed that combining all the features (both with **SVM** and **KNN**) is resulting in better performance compared to others.

Features	Recognition rate
HSV	43.0%
SIFT internal	55.1%
SIFT boundary	32.0%
HOG	49.6%
HSV + SIFT int	66.4%
HSV + SIFT bdy	57.0%
HSV + HOG	62.1%
SIFT int + SIFT bdy	58.6%
SIFT int + HOG	66.4%
SIFT bdy + HOG	55.3%
HSV + SIFT int + HOG	71.8%
HSV + SIFT int + SIFT bd + HOG	72.8%

Bilinear pyramid network for flower species categorization - Year 2020 by Cheng Pang · Wenhao Wang · Rushi Lan · Zhuo Shi · Xiaonan Luo

- They proposed **Bilinear Pyramid Network** (BPN) for flower categorization. The proposed BPN is built on the **VGG-16** architecture. When passing an image through a **CNN**, feature maps keep decreasing in their sizes, forming a feature pyramid.
- They have used 2 CNN's of the bilinear model. One for feature detection and another for feature extraction.
- They first introduced up-sampling layers and bilinear pooling layers to the last three convolutional layers conv5 1, conv5 2 and conv5 3, obtaining the bilinear product of their features. Then the product is fed through the fully connected layers f c 7 and f c 8, where f c 8 is the predefined VGG-16 layer and f c 7 is newly added for enhancing the non-linear property of the network.
- The training process follows the standard stochastic gradient descent (**SGD**) with a learning rate of 0.01, and will be terminated when reach the maximum iteration of 100000.
- These two species of flowers (ColtsFoot and Dandelion) share similar delicate long-strip petals. Their difference only lies in the stamen. As a result, it can be confused to differentiate some over-exposure and low-resolution samples with ambiguous details.
- They are able to achieve a top-1 precision of **99.1%** using 3 convolutional layers in the feature pyramid are used.

[Union-net: A deep neural network model adapted to small data sets - Year 2020 by Qingfang He , Guang Cheng , Zhiying Lin]

- In this paper they have proposed a Union-net model which is based on the convolutional neural network (**CNN**) technology. Convolutional network units with different combinations of the same input form a Union module. Each Union module is a convolutional layer.
- A "3-layer" neural network is formed through the serial input and output between the three modules. The outputs of the three Union modules are fused and added as the input of the last convolutional layer, thus forming a complex network with a 4-layer network structure
- Union-net uses a maximum pooling layer in the Union 1 module. Max pooling can extract effective features, reduce dimensions and parameters and remove noise.
- After the initial input data of the model is processed by Union1's internal convolution structure, the output features are pooled and filtered to provide more effective data for subsequent processing.
- The model structure adopts 4-3-1 combination.
- They are able to achieve an accuracy of **87%** with 10-fold cross-validation.

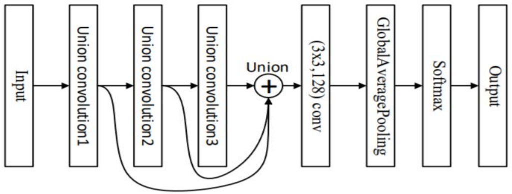


Fig3. Union-net model simplified with Union convolution

[Classification of flower image based on attention mechanism and multi-loss attention network : Year 2021 by Mei Zhang, Huihui Su, Jinghua Wen]

- This paper proposes the multi-loss spatial attention network (**MLSAN**), the multi-loss channel attention network (**MLCA**) and the multi-loss multi-attention network (**MLMAN**) three network models, with Xception as the basic network, channel attention and spatial attention are added to Xception to improve Xception's ability to locate and extract features of flower image regions.
- The multi-attention loss neural network designed in this paper works while introducing attention mechanisms, the neural network in this paper combines the multiple loss functions of Triplet Loss and Softmax Loss to jointly optimize the network, thereby further improving the accuracy of flower image classification
- They are able to achieve a maximum accuracy of **98.03%** Using MLCSCAN method.

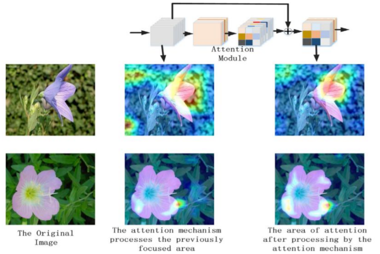


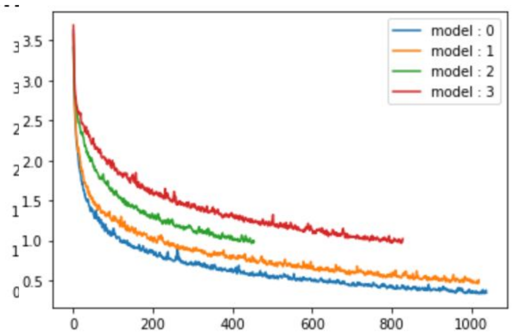
Fig. 17. Visualization results of class activation maps before and after the attention mechanism is added to the Oxford 102 flowers data set.

Method	Oxford 17 flowers
Xception	97.31%
MLSAN	97.70%
MLCAN	97.81%
MLCSAN	98.03%

Result Analysis

- In this project 4 models are considered by combination of features.
  - Model 1 : hsv and siftint
  - Model 2 : hsv and siftbdy
  - Model 3 : hog and siftint
  - Model 4 : hog and siftbdy
- 4 models are trained individually and then for a particular pattern, each one of the model is predicted and out of the predicted classes by different models, the class which has highest number of predictions is considered as final predicted class for a particular pattern. Ex : for pattern 1
  - Model 1 predicts class 3
  - Model 2 predicts class 4
  - Model 3 predicts class 16
  - Model 4 predicts class 3
- So class 3 has the highest number of predictions and it is considered as the final predicted class for that pattern
- Machine Learning Algorithms are run using k-fold cross validation with k = 5 and k = 10 as train size vary when k = 5 and k = 10.

- **Logistic Regression :**
    - For 5-fold it's able to achieve an accuracy of **80.41%**
    - For 10-fold it's able to achieve an accuracy of **80.07%**
    - It's performing best compared to all other algorithms that are evaluated.
- **Single Layer Perceptron :**
    - For 5-fold it's able to achieve an accuracy of **77.86%**
    - For 10-fold it's able to achieve an accuracy of **78.97%**
    - It's slightly closer to the Logistic Regression Algorithm accuracy with 2-3% variance.
    - SLP without k-fold (train : 70% ; test : 30 %) :
        - Model 0 Score : 79.41
        - Model 1 Score : 71.81
        - Model 2 Score : 62.74
        - Model 3 Score : 57.35
        - Accuracy Percentage : 79.65



- **Perceptron Learning Algorithm :**
    - For 5-fold it's able to achieve an accuracy of **71.69%**
    - For 10-fold it's able to achieve an accuracy of **70.44%**

- **Multi Layer Perceptron :**
  - For 5-fold it’s able to achieve an accuracy of **71.54%** with 1 hidden layer of 2000 hidden neurons and max iterations of 1500.
  - For 10-fold it’s able to achieve an accuracy of **53.60%** with 1 hidden layer of 100 hidden neurons and max iterations of 1500.
  - It can be observed that as the number of hidden neurons increases, the more linearity the trained dataset will tend to and increases the accuracy. Which is described by the **cover’s** theorem.

- Accuracy Comparison :
  - LogisticRegression > SLP > MLP > PLA
  - The accuracy of the MLP can be increased by further increasing the number of hidden neurons.
  - It’s also can be observed that the accuracy is almost similar in the respective algorithm in case of 5-fold and 10-fold cross validation.

	LogisticRegression		SLP		PLA	
HSV	57.57	57.79	54.19	57.42	42.72	42.05
SIFT internal	66.98	67.27	60.07	61.47	52.57	48.67
SIFT boundary	58.01	59.11	45.36	47.64	36.76	33.38
HOG	49.55	51.17	46.91	46.91	33.38	35.36
HSV + SIFT int	79.11	80.29	77.72	78.23	70.88	65.36
HSV + SIFT bdy	73.16	74.11	71.83	70.51	57.94	62.42
HSV + HOG	68.82	70.0	66.10	68.30	58.97	59.04
SIFT int + SIFT bdy	61.02	60.36	58.30	62.20	46.02	50.44
SIFT int + HOG	69.85	70.22	68.16	68.23	59.48	59.70
SIFT bdy + HOG	69.55	70.80	62.79	63.97	47.79	46.61
HSV + SIFT int + HOG	80.44	81.39	79.41	80.58	73.60	72.86
HSV + SIFT int + SIFT bdy + HOG	81.32	83.16	80.44	80.22	71.17	72.27

- Instead of considering combination of features as models, here we combined various combinations of features and experimented using ML algorithms.
- Observed Results are shown above. Each algorithm has 2 values, first value represents the accuracy for 5-fold and second value is for 10-fold.
- It can be observed that combining almost all the features is giving good accuracy compared to single feature ones.

- Another method which is different from those that were affirmed earlier.
- Converting Image to Gray Scale and resize to (124,124) then to single dimensional in which each pattern contains 15376 features.

- With 5-fold this approach is able to achieve an accuracy of **23.30%**
- Example image of Gray scale :

THANK YOU