**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**KHOA KHOA HỌC MÁY TÍNH**

**-------------------------**



**BÁO CÁO ĐỒ ÁN MÔN NHẬP MÔN THỊ GIÁC MÁY TÍNH**

***ĐỀ TÀI***

**OBJECT CLASSIFICATION**

**Thành viên nhóm:**

1. **Đặng Vũ Minh Hiếu 15520218**
2. **Võ Quốc Hưng 15520289**
3. Giới thiệu để tài:
   1. Giới thiệu bài toán:

Cho một ảnh cho trước, hãy gán nhãn (label) cho ảnh đó

vd :

* Nhãn : dog



* Nhãn : airplane
  1. Ứng dụng
* Phân loại ảnh có vi phạm giao thông hay không
* Phân loại trái cây hư
* Phân loại giới tính

1. Phương pháp giải quyết

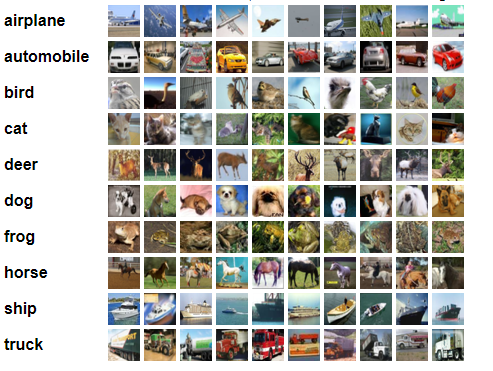
Về cơ bản, object classification là một bài toán phân lớp với nhiều lớp khác nhau (multiclass). Phương pháp phổ biến nhất được dùng để giải bài toán toán phân hai hoặc nhiều lớp là máy học (machine learning)

Có nhiều phương pháp máy học áp dụng cho bài toán phân lớp như : Naive Bayes, SVM,… và chúng đều có chung một đặc điểm là cần một tập dữ liệu cho trước để máy tính có thể “học” và đưa ra dự đoán một nhãn cụ thể cho một ảnh

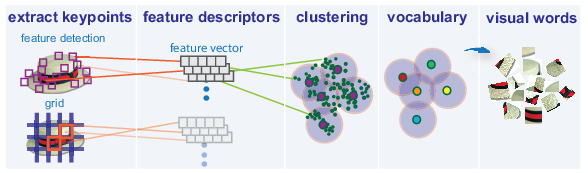
* 1. Mô tả tập dữ liệu.

Ở đây chúng ta sẽ sử dụng tập dữ liệu CIFAR-10:

Tập dữ liệu CIFAR-10 chứ 60000 tấm ảnh màu có kích thước 32x32 được phân bố đều cho 10 lớp, mỗi lớp 6000 tấm ảnh như sau:



* 1. Rút trích đặc trưng

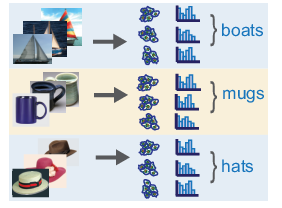
Với dữ liệu thô như trên chỉ bao gồm các ảnh với các giá trị pixel thường không mang nhiều ý nghĩa, hoặc có lợi cho phương pháp máy học. Vì thế ta cần phải tìm cách biến đổi dữ liệu để phục vụ cho việc huấn luyện bộ phân lớp. Một phương pháp phổ biến đó là sử dụng Bag of Words (túi đựng từ) với cách làm như sau:

* Đầu tiên ta rút trích các đặc trưng cục bộ của các ảnh sử dụng thuật toán SIFT để thu được các đặc trưng cục bộ của các ảnh, được biểu diễn qua một tập hợp feature descriptors dưới dạng các vector có kích thước 128
* Sau đó ta tiến hành gom nhóm các feature vector bằng cách sử dụng phương pháp K-mean clustering ta sẽ thu được các điểm trung tâm của các nhóm thì tập các điểm nãy sẽ là túi đựng từ mà ta cần tìm
  1. Các phương pháp máy học

Túi đựng từ là công cụ để giúp ta lấy ra được các feature vector cho tập dữ liệu train của các phương pháp máy học, ta cần phải làm một số thao tác để có thể rút ra được các feature vector cho từng ảnh:



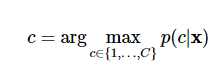
* Đầu tiên với một ảnh ta tiến hành rút trích đặc trưng cục bộ, sau đó phân phối các feature vector và các cluster (hay word trong túi đựng từ). Với mỗi một cluster ta sẽ đếm số lượng point thuộc cluster đó, tương đương với giá trị ở vị trí tương ứng trong feature vector. Kết quả ta thu được một feature vector với số chiều bằng số lượng word trong túi đựng từ
* Lặp lại thao tác trên với tất cả các ảnh trong tập train.

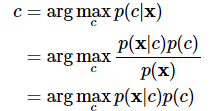


* Sau khi kết thúc ta thu được tập dữ liệu đã được xử lý để có thể áp dụng các phương pháp máy học
  + 1. Phương pháp Naive Bayes

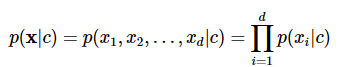
Xét bài toán classification với C classes 1,2,…,C. Giả sử có một điểm dữ liệu x∈Rd. Hãy tính xác suất để điểm dữ liệu này rơi vào class c. Nói cách khác, hãy tính:



Biểu thức này, nếu tính được, sẽ giúp chúng ta xác định được xác suất để điểm dữ liệu rơi vào mỗi class. Từ đó có thể giúp xác định class của điểm dữ liệu đó bằng cách chọn ra class có xác suất cao nhất:

Tuy vậy biểu thức trên hơi khó để tính nên ta thường sử dụng như sau:

Trong đó

P(x|c) = Phân phối các điểm dữ liệu trong class c được tính như sau

P(c )= Xác suất để một điểm rơi vào class c. Thường được tính bằng tỉ lệ số điểm dữ liệu trong tập training rơi vào class này chia cho tổng số lượng dữ liệu trong tập traing

Vậy đối với bài toán Object Classification thì ta áp dụng Naive Bayes như sau:

* Xét một ảnh cho trước ta rút trích đặc trưng của ảnh đó thu được feature vector x = {x1, x2, x3,…., xn}
* Xét lần lượt các class. Tính p(x|c) và p(c ) theo công thức trên
* Kết quả phân lớp chính là class cho ra p(x|c)p(c ) lớn nhất

1. Kết quả thực nghiệm
   1. Tổ chức dữ liệu

Bộ dữ liệu CIFAR-10 có 60000 tấm ảnh và 10 lớp, mỗi lớp 1000 tấm ảnh. Trong đó:

* Dữ liệu để train là 50000 ảnh
* Dữ liệu để test là 10000 ảnh. Trong bộ dữ liệu test thì bao gồm 1000 ảnh ngẫu nhiễn ở mỗi lớp. Các ảnh khác sẽ nằm trong bộ dữ liệu train.
  1. Chương trình

Chương trình được viết bằng ngôn ngữ Python 2 với các thư viện bắt buộc sau:

* tKinter
* OpenCv
* Sklearn
* Pytouch
* Và các thư viện khác trong file requirement.txt
  1. Kết quả

|  |  |
| --- | --- |
| Phương pháp | Độ chính xác (%) |
| Naive Bayes | 27.73 |
| SVM Polynomial | 28.34 |
| SVM RBF | 20.31 |
| ANN | 29.67 |
| Softmax Regression | 29.26 |

Biểu đồ độ chính xác của từng classifier:

Nhận xét: Các classifier đều cho kết quả khá tương đồng nhau, ngoại trừ SVM RBF. Mặt khác độ chính xác của các phương pháp này còn thấp đối với tập dữ liệu đã cho

1. Kết luận

Nhìn chung với tập dữ liệu CIFAR-10, độ chính xác khi phân lớp sử dụng các phương pháp máy học và túi đựng từ là không cao, một phần là do tập dữ liệu tương đối khó (ảnh nhỏ, object một class đa dạng). Để cải thiện hiệu suất, ta nên sử dụng các phương pháp khác như Deep Feature, Deep Learning

1. Tài liệu tham khảo

-<https://www.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html>

- International Journal of Computer Science and Electronics Engineering (IJCSEE) Volume 4, Issue 3 (2016) ISSN 2320–4028 (Online). Image Classification Using Naïve Bayes Classifier Dong-Chul Park

1. Phân công công việc nhóm

|  |  |  |
| --- | --- | --- |
| MSSV | Họ và tên | Công việc thực hiện |
| 15520218 | Đặng Vũ Minh Hiếu | Xây dựng GUI, Naive Bayes, đánh giá các classifier |
| 15502289 | Võ Quốc Hưng | Xây dưng túi đựng từ classifier SVM, ANN, Softmax |