

BÁO CÁO ĐỒ ÁN CUỐI KỲ MÔN MÁY HỌC

Lớp CS114.K21.KHTN

Giảng viên hướng dẫn: Thầy Lê Đình Duy

Thầy Phạm Nguyễn Trường An

Sinh viên thực hiện: Nguyễn Trần Trung - 18521555

I. Mô tả bài toán

Bài toán: Phân loại các tòa nhà trong khuôn viên UIT

- Input: Một bức ảnh các tòa nhà trong khuôn viên UIT
- Output: Đó là tòa nhà nào trong UIT

Kết quả đạt được:

- Đã xây dựng được nhiều model với các phương pháp lấy đặc trưng khác nhau
- Model tốt nhất đạt được kết quả hơn 60%

II. Mô tả bộ dữ liệu

Về bộ dữ liệu: Bộ dữ liệu được em cùng 3 bạn khác là Nguyễn Khánh Toàn, Trần Đình Khang (18520072), Võ Quốc An cùng thu thập.
+ Bộ dữ liệu gồm hơn 1361 ảnh. Trong đó có 2 lần chụp vào 2 ngày khác nhau. Ngày thứ nhất chụp 1029 bức. Ngày thứ 2 chụp 331 bức.

+ Ảnh được chia vào 5 lớp khác nhau gồm có : “nhà B”, “nhà C”, “Canteen”, “nhà D”, “nhà E”

+ Số lượng tương ứng với từng lớp là:

Nhà B: 320

Nhà C: 280

Canteen: 217

Nhà D: 233

Nhà E: 306

III. Tiền xử lý dữ liệu và rút trích đặc trưng

- Tiền xử lý dữ liệu: resize ảnh, chuyển ảnh sang ảnh màu RGB

Trích xuất đặc trưng: Về xử lý ảnh thì phần này nằm ngoài phạm vi của môn Máy học, tuy nhiên em có tìm hiểu một số phương pháp để rút trích đặc trưng như sử dụng thư viện OpenCV, hoặc skimage. Trong đó có những phương pháp như Visual bag of word, Histogram of Oriented gradient, SUFT,... Trong đồ án môn học, em chọn sử dụng thư viện skimage và phương pháp rút trích đặc trưng Histogram of Oriented gradient giải quyết bài toán của em. Ngoài ra em còn thử chuyển bức ảnh thành array để làm input đầu vào cho model để có được nhiều cái nhìn khác nhau.

Em phân chia bộ dữ liệu thành 2 phần:

Phần 1 gồm 1026 ảnh chụp vào ngày đầu tiên. Đó sẽ là gồm training set và validation set. Tỷ lệ validation set là 30% của bộ dữ liệu.

Phần 2 gồm 331 ảnh chụp vào ngày thứ 2. Đây sẽ được coi là unseen data để test lại mô hình đã được huấn luyện bởi bộ dữ liệu bên trên

Về phương pháp rút trích đặc trưng HOG, vector đặc trưng sẽ được

tính như sau:

- $\text{pixels_per_cell}(x,y)$: ảnh được chia ra thành các cell có số pixel là $x*y$
- $\text{cells_per_block}(x,y)$: mỗi block chứa $x*y$ ô cells
- orientations : số hướng trong mỗi cell

giả sử chúng ta có ảnh là $64 * 128$ ($w * h$), orientations = 9,
 $\text{pixels_per_cell}(8,8)$, $\text{cells_per_block}(3,3)$

-> thì với số cells có được là $8 * 16$

mỗi block sẽ bao gồm số cell = cells_per_block . Ở ví dụ đây là mỗi block gồm có $3*3$ cell = 9 cell.

Các block được xét sẽ gồm có $3*3$ cell và xét cả các block có chồng lên nhau.

-> số block là $8 - (3 - 1) * 16 - (3 - 1) = 6 * 14$

với mỗi block, ta tính được ta tính được là $9 * 3 * 3 = 81$ vector thành phần.

-> chiều dài của features vector : $81 * 14 * 6 = 6804$

Đối với phương pháp rút trích đặc trưng HOG sẽ nhận được một vector đặc trưng và một ảnh biểu diễn vector đặc trưng. Từ ảnh biểu diễn vector đặc trưng em sẽ tăng cường cường độ của những đặc trưng có trong ảnh và chuyển ảnh vừa tăng cường thành mảng để đưa vào trong thuật toán huấn luyện.

III. Mô hình thuật toán máy học

Trong Classifications, có nhiều thuật toán phân loại khác nhau như SVM, Decision Tree, ...

Trong đồ án, em chọn thuật toán SVM, SGD, Decision Tree để giải quyết bài toán

Đối với mỗi thuật toán, em sử dụng 3 phương pháp rút trích đặc trưng khác nhau:

1. Kết quả của phương pháp sử dụng ảnh gốc chuyển sang vector đặc trưng trên bộ dữ liệu huấn luyện.

```
0.487012987012987
<class 'sklearn.tree._classes.DecisionTreeClassifier'>
precision    recall  f1-score   support

      B       0.62       0.60       0.61         78
      C       0.39       0.45       0.42         69
      CT       0.27       0.25       0.26         44
      D       0.66       0.55       0.60         53
      E       0.48       0.50       0.49         64

 accuracy          0.49         308
macro avg       0.48       0.47       0.47         308
weighted avg    0.49       0.49       0.49         308

0.8506493506493507
<class 'sklearn.svm._classes.SVC'>
precision    recall  f1-score   support

      B       0.92       0.87       0.89         78
      C       0.76       0.91       0.83         69
      CT       0.80       0.80       0.80         44
      D       0.94       0.91       0.92         53
      E       0.86       0.75       0.80         64

 accuracy          0.85         308
macro avg       0.85       0.85       0.85         308
weighted avg    0.86       0.85       0.85         308

0.788961038961039
<class 'sklearn.linear_model._stochastic_gradient.SGDClassifier'>
precision    recall  f1-score   support

      B       0.79       0.88       0.84         78
      C       0.75       0.80       0.77         69
      CT       0.79       0.61       0.69         44
      D       0.91       0.81       0.86         53
      E       0.73       0.77       0.75         64

 accuracy          0.79         308
macro avg       0.80       0.77       0.78         308
weighted avg    0.79       0.79       0.79         308
```

2. Kết quả của phương pháp rút trích đặc bằng HOG trên tập dữ liệu huấn luyện

```

<class 'sklearn.svm._classes.SVC'>
precision    recall  f1-score   support

      B      0.79      0.77      0.78        78
      C      0.81      0.86      0.83        69
      CT      0.78      0.70      0.74        44
      D      0.96      0.85      0.90        53
      E      0.75      0.84      0.79        64

 accuracy      0.81
 macro avg      0.81
 weighted avg    0.81

0.8084415584415584

<class 'sklearn.linear_model._stochastic_gradient.SGD'>
precision    recall  f1-score   support

      B      0.46      0.96      0.62        78
      C      0.96      0.35      0.51        69
      CT      0.79      0.77      0.78        44
      D      0.97      0.62      0.76        53
      E      0.98      0.64      0.77        64

 accuracy      0.67
 macro avg      0.83
 weighted avg    0.81

0.672077922077922

<class 'sklearn.tree._classes.DecisionTreeClassifier'>
precision    recall  f1-score   support

      B      0.49      0.47      0.48        78
      C      0.53      0.43      0.48        69
      CT      0.20      0.20      0.20        44
      D      0.47      0.53      0.50        53
      E      0.33      0.38      0.35        64

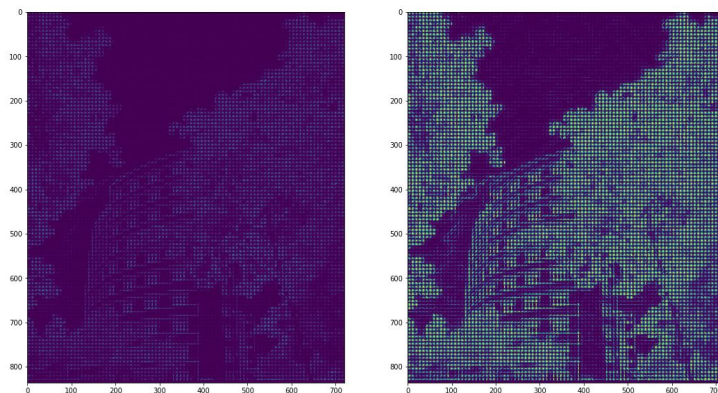
 accuracy      0.42
 macro avg      0.41
 weighted avg    0.42

0.4155844155844156

```

3. Kết quả của phương pháp sử dụng ảnh HOG và tăng cường cường độ đặc trưng ảnh trên bộ dữ liệu huấn luyện

Ảnh trước và sau khi tăng cường độ đặc trưng ảnh:



Kết quả:

```
0.8521400778210116
precision    recall  f1-score   support

   B         0.91     0.81     0.85        62
   C         0.76     0.93     0.84        61
  CT         0.82     0.73     0.77        37
   D         1.00     0.93     0.96        41
   E         0.84     0.84     0.84        56

 accuracy          0.85        257
 macro avg         0.87     0.85     0.85        257
weighted avg         0.86     0.85     0.85        257
```

IV. Kết quả đạt được trên dữ liệu kiểm thử

1. Kết quả của phương pháp sử dụng ảnh gốc chuyển sang vector đặc trưng.

```
<class 'sklearn.svm._classes.SVC'>
0.4169184290030212
precision    recall  f1-score   support

   B         0.35     0.93     0.51        58
   C         0.00     0.00     0.00        67
  CT         0.83     0.22     0.35        67
   D         0.84     0.87     0.86        79
   E         0.00     0.00     0.00        60

 accuracy          0.42        331
 macro avg         0.41     0.41     0.34        331
weighted avg         0.43     0.42     0.37        331
```

2. Kết quả của phương pháp rút trích đặc bằng HOG

```
ss 'sklearn.svm._classes.SVC'>
09969788519638
precision    recall  f1-score   support

   B         0.50     0.47     0.48        58
   C         0.60     0.22     0.33        67
  CT         0.83     0.52     0.64        67
   D         1.00     0.90     0.95        79
   E         0.29     0.68     0.41        60

 accuracy          0.57        331
acro avg         0.65     0.56     0.56        331
```

3. Kết quả của phương pháp sử dụng ảnh HOG và tăng cường cường độ đặc trưng ảnh.

```
<class 'sklearn.svm._classes.SVC'>  
0.6465256797583081
```

	precision	recall	f1-score	support
B	0.79	0.72	0.76	58
C	0.42	0.28	0.34	67
CT	0.64	0.76	0.69	67
D	0.98	0.77	0.87	79
E	0.45	0.68	0.54	60
accuracy			0.65	331
macro avg	0.66	0.64	0.64	331
weighted avg	0.67	0.65	0.65	331

V. Kết luận

Mô hình vẫn chưa thực sự tốt (mô hình tốt nhất khoảng 64%). Những nguyên nhân dẫn đến là:

- + Bộ dữ liệu còn quá ít. (khoảng 150 ảnh cho mỗi lớp)
- + Tiền xử lý dữ liệu và rút trích đặc trưng còn hạn chế về kiến thức
- + Trong quá trình thu thập dữ liệu thì những tòa nhà rất lớn, không thể chụp vào một khung ảnh nên phải chụp từng góc. Điều này đã dẫn đến việc giống nhau giữa một số tòa nhà nên dẫn đến việc giảm độ chính xác của mô hình.