

PHÂN LOẠI CÁC TÒA NHÀ TRONG KHUÔN VIÊN UIT

NGUYỄN TRẦN TRUNG - 18521555 - CS114.K21.KHTN

Link github: github.com/shaking54/CS114.K21.KHTN

Tóm tắt

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%

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 trong

Nhà B: 320

Nhà C: 280

Canteen: 217

Nhà D: 233

Nhà E: 306

Tiền xử lý dữ liệu & rút trích đặc trưng

Về tiền xử lý dữ liệu:

- Chuyển ảnh về ảnh màu RGB
- Resize ảnh thành (256,256)

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.

Phân chia dataset

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

Chọn mô hình thuật toán máy học

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

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

Về classifications nhiều nhãn: Có 2 loại đó là OnevsOne và OnevsTheRest

+ Về OneVsOne thì tại một thời điểm, class được bỏ phiếu nhiều nhất thì sẽ được chọn.

=> sẽ có $n_classes * (n_classes - 1) / 2$ classifiers

+ Đối với OnevsTheRest thì mỗi class sẽ có 1 classifiers

Chọn mô hình thuật toán máy học

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.

```
0.8506493506493507
<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

Mô hình SVC

```
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

Mô hình SGD

```
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

Mô hình Decision Tree

Chọn mô hình thuật toán máy học

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         308
 macro avg         0.82         0.80      0.81         308
 weighted avg      0.81         0.81      0.81         308

0.8084415584415584
```

Mô hình SVC

```
<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         308
 macro avg         0.83         0.67      0.69         308
 weighted avg      0.81         0.67      0.67         308

0.672077922077922
```

Mô hình SGD

```
<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

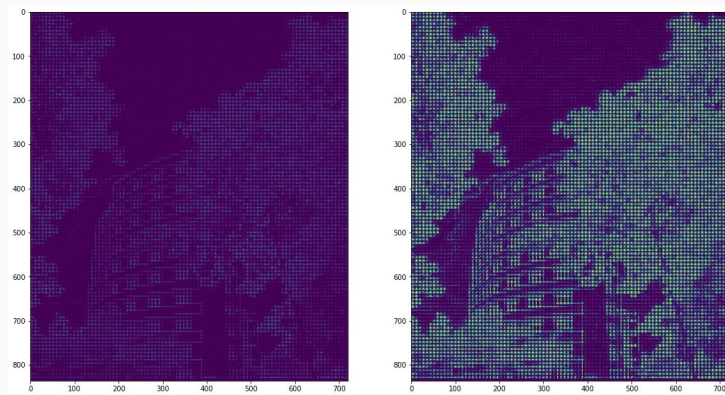
 accuracy          0.42         308
 macro avg         0.41         0.40      0.40         308
 weighted avg      0.42         0.42      0.42         308

0.4155844155844156
```

Mô hình Decision Tree

Chọn mô hình thuật toán máy học

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



Ảnh trước và sau khi tăng cường cường độ

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

Sử dụng model cho Unseen data

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

Mô hình SVC

```
<class 'sklearn.linear_model._stochastic_gradient.SGDC'>  
0.38368580060422963
```

	precision	recall	f1-score	support
B	0.28	0.95	0.43	58
C	0.00	0.00	0.00	67
CT	0.37	0.15	0.21	67
D	0.90	0.68	0.78	79
E	0.26	0.13	0.18	60
accuracy			0.38	331
macro avg	0.36	0.38	0.32	331
weighted avg	0.39	0.38	0.34	331

Mô hình SGD

```
<class 'sklearn.tree._classes.DecisionTreeClassifier'>  
0.32326283987915405
```

	precision	recall	f1-score	support
B	0.44	0.69	0.54	58
C	0.24	0.21	0.22	67
CT	0.31	0.31	0.31	67
D	0.81	0.32	0.45	79
E	0.08	0.12	0.10	60
accuracy			0.32	331
macro avg	0.38	0.33	0.33	331
weighted avg	0.40	0.32	0.33	331

Mô hình Decision Tree

Sử dụng model cho Unseen data

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

Mô hình SVC

0.525679758308157

	precision	recall	f1-score	support
B	0.29	0.78	0.43	58
C	1.00	0.04	0.09	67
CT	0.82	0.82	0.82	67
D	1.00	0.65	0.78	79
E	0.35	0.33	0.34	60
accuracy			0.53	331
macro avg	0.69	0.52	0.49	331
weighted avg	0.72	0.53	0.51	331

Mô hình SGD

0.27794561933534745

	precision	recall	f1-score	support
B	0.27	0.34	0.30	58
C	0.31	0.21	0.25	67
CT	0.18	0.13	0.15	67
D	0.49	0.29	0.37	79
E	0.23	0.43	0.30	60
accuracy			0.28	331
macro avg	0.30	0.28	0.27	331
weighted avg	0.30	0.28	0.28	331

Mô hình Decision Tree

Sử dụng model cho Unseen data

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

Kết quả và đánh giá

Các model sử dụng đã bị overfitting. Nguyên nhân dẫn đến bởi các khó khăn sau:

- Bộ dữ liệu còn quá ít (khoảng 1400 ảnh cho 5 class) chưa có được sự tổng quát cho mô hình.
- Phương pháp trích xuất đặc trưng còn hạn chế, chưa loại bỏ được những đối tượng không mong muốn trong hình
- Hạn chế về kỹ năng thu thập dữ liệu. Các nhà B, E, C khá to, không thể chụp toàn bộ một tòa nhà trong một khung hình. Điều này dẫn đến sự có sự tương đồng giữa các tòa nhà nên đã gây nên sự nhầm lẫn cho mô hình dự đoán