**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

****

**BÁO CÁO KỸ THUẬT ĐỒ ÁN**

**MÔN NHẬN DẠNG THỊ GIÁC VÀ ỨNG DỤNG**

**TP. HỒ CHÍ MINH, 07/2017**

**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

****

**CH1501020 – DƯƠNG THÁI BẢO**

**BÁO CÁO KỸ THUẬT ĐỒ ÁN**

**MÔN NHẬN DẠNG THỊ GIÁC VÀ ỨNG DỤNG**

**GIẢNG VIÊN GIẢNG DẠY**

**TS. LÊ ĐÌNH DUY**

**TS. NGUYỄN TẤN TRẦN MINH KHANG**

**TP. HỒ CHÍ MINH, 07/2017**

**LỜI CẢM ƠN**

Báo cáo đồ án môn học là một trung những đánh giá thiết thực nhất về khả năng tiếp thu cũng như khả năng hiện thực hoá các nội dung đã được truyền đạt trong quá trình học tập, vì vậy việc hoàn thành báo cáo đồ án môn học thực sự cần đến nhiều sự giúp đỡ từ các yếu tố bên ngoài để có thể hoàn thành một cách tốt nhất.

Trong quá trình thực hiện học tập và thực hiện báo cáo đồ án môn học em đã được sự giúp đỡ, hướng dẫn, hỗ trợ và động viên từ gia đình, từ quý thầy cô cùng các bạn. Nhờ đó mà em đã hoàn thành được báo cáo.

Em xin chân thành cảm ơn TS. Lê Đình Duy và TS. Nguyễn Tấn Trần Minh Khang đã trực tiếp giảng dạy. Trong quá trình học tập, và hoàn thành báo cáo thầy đã tận tình truyền đạt kiến thức cũng như hướng dẫn thực hiện báo cáo này.

|  |  |
| --- | --- |
|  | Học viên  Dương Thái Bảo |

|  |  |
| --- | --- |
|  |  |

# NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

.......................................................................................................................................

# DANH MỤC HÌNH ẢNH

[Hình 1.1 Cấu trúc thư mục 5](#_Toc488091311)

[Hình 1.2 Cấu trúc thư mục sau khi tải về. 6](#_Toc488091312)

[Hình 2.1 Cell size = [4 4] 7](#_Toc488091313)

[Hình 2.2 Cell size =[8 8] 8](#_Toc488091314)

[Hình 3.1 Bag of visual words. 8](#_Toc488091315)

[Hình 3.2 Quá trình xây dựng bag of word 9](#_Toc488091316)

[Hình 3.3 Xây dựng bag of feature trong matlab 9](#_Toc488091317)

[Hình 3.4 Bag Of Feature với các tham số mặc định 10](#_Toc488091318)

[Hình 3.5 Bag of Feature với CustomExtractor 10](#_Toc488091319)

[Hình 3.6 Áp dụng tham số feature layer = ‘conv4’ 11](#_Toc488091320)

[Hình 3.7 Áp dụng feature layer = ‘fc8’ 12](#_Toc488091321)

[Hình 3.8 Áp dụng feature layer = ‘fc7’ 12](#_Toc488091322)

# MỤC LỤC

LỜI CẢM ƠN

[NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN 2](#_Toc488036369)

[DANH MỤC HÌNH ẢNH 3](#_Toc488036370)

[MỤC LỤC 4](#_Toc488036371)

[CHƯƠNG I TỔNG QUAN VỀ ĐỒ ÁN 5](#_Toc488036372)

[1. Giới thiệu: 5](#_Toc488036373)

[2. Tổ chức của chương trình: 5](#_Toc488036374)

[3. Cài đặt: 5](#_Toc488036375)

[CHƯƠNG II XÂY DỰNG ỨNG DỤNG 7](#_Toc488036376)

[1. Phương pháp 1: 7](#_Toc488036377)

[a. Các bước thực hiện: 7](#_Toc488036378)

[b. Kết quả nhận xét 7](#_Toc488036379)

[2. Phương pháp 2: Bag Of Word 8](#_Toc488036380)

[a. Các bước thực hiện: 9](#_Toc488036381)

[b. Kết quả nhận xét: 10](#_Toc488036382)

[3. Phương pháp 3: Deep Learning 10](#_Toc488036383)

[a. Các bước thực hiện: 10](#_Toc488036384)

[b. Kết quả nhận xét: 11](#_Toc488036385)

[4. Đánh giá: 12](#_Toc488036386)

[KẾT LUẬN 13](#_Toc488036387)

[HƯỚNG PHÁT TRIỂN 13](#_Toc488036388)

[DANH MỤC TÀI LIỆU THAM KHẢO 14](#_Toc488036389)

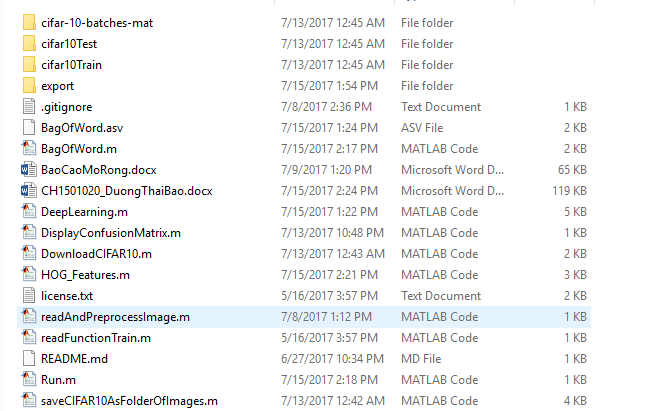
# CHƯƠNG I TỔNG QUAN VỀ ĐỒ ÁN

## Giới thiệu:

Image classification là gán nhãn cho đối tượng thuộc một class. Đồ án sử dụng một số phương pháp phân lớp lên tập dữ liệu. Đồ án sẽ tìm hiểu về bài toán phân loại ảnh (image classification) dựa vào bộ dữ liệu CIFAR10 [5].

## Tổ chức của chương trình:

Chương trình có cấu trúc như sau:



**Hình 1.1 Cấu trúc thư mục**

Cácthư mục: **cifar-10-batches-mat**, **cifar10Test**, **cifar10Train** là thư mục chứa dữ liệu train và test.

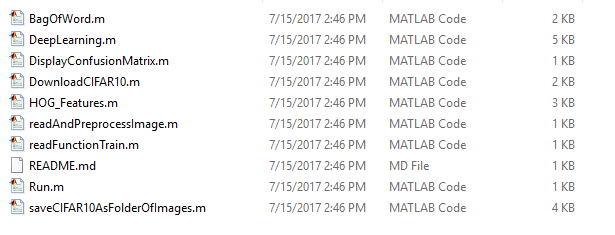
Thư mục **export** là nơi lưu trữ các biến variable, có định dạng là .mat

**Run.m** file là để thực thi và hiển thị kết quả.

Các file DeepLearning.m, HOG\_Features.m, BagOfWord.m là các hàm xử lý sử dụng lần lượt các feature và phương pháp: Deep Learning, HOGFeatures và SVM, Bag Of Word (BoW).

## Cài đặt:

Bước 1: clone soure từ github, url: <https://github.com/forverbao990/vra-project.git>



**Hình 1.2 Cấu trúc thư mục sau khi tải về.**

Bước 2: Execute ***DownloadCIFAR10.m*** để chuẩn bị dữ liệu train và test.

# CHƯƠNG II XÂY DỰNG ỨNG DỤNG

## Phương pháp 1:

Mục đích của chương này là sử dụng HOG features và phương pháp SVM (Support Vector Machine) trong bài toán phân loại ảnh.

Phương pháp phân loại ảnh dùng HOG Features được sử dụng nhiều trong nhận dạng ký tự quang học (Optical Character Recognition), áp dụng nhiều trong bài toán nhận dạng chữ số viết tay.

Tổng quát của phương pháp này là:

* Rút trích đặc trưng (feature) từ tập dữ liệu train và tập dữ liệu test
* Phân loại đặc trưng dùng SVM (Support Vector Machine)

## Các bước thực hiện:

Bước 1: Phân vùng bộ dữ liệu thành tập huấn luyện và tập kiểm tra, gồm các loại nhãn: **Deer**, **Dog**, **Frog**, **Cat** và **Ship.** Sử dùng hàm *imageDatastore* trong matlab.

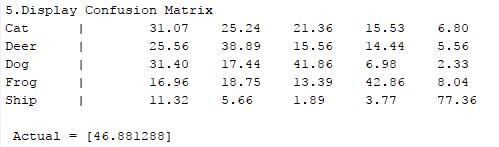
Bước 2**:** Trích xuất HOG feature trong bộ dữ liệu train và HOG feature trong bộ dữ liệu test.

Bước 3: Tiếp theo, xây dựng model từ tập dữ liệu train dùng phương pháp SVM. Sử dụng hàm **fitcecoc** trong matlab để xây dựng model.

Bước 4: Hiển thị ma trận và đánh giá độ chính xác.

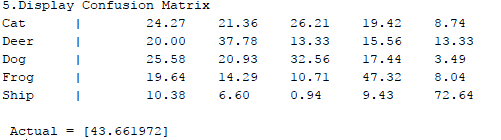
## Kết quả nhận xét

Cell size = [4 4], kết quả:



**Hình 2.1 Cell size = [4 4]**

Cell size = [8 8], kết quả:



**Hình 2.2 Cell size =[8 8]**

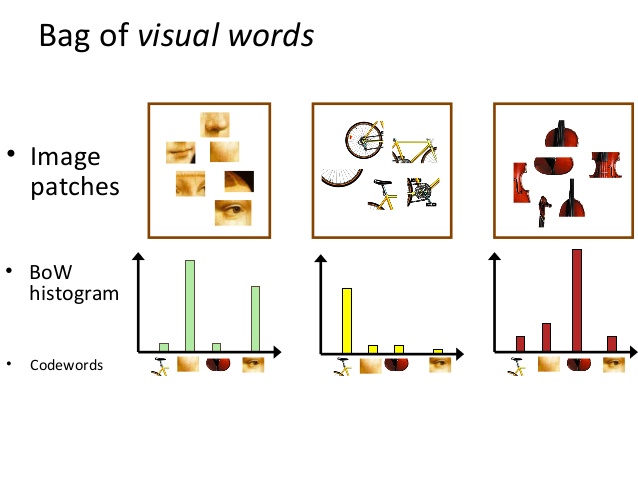
## Phương pháp 2:

Bag of Features có dạng như sau:

Trong đó:

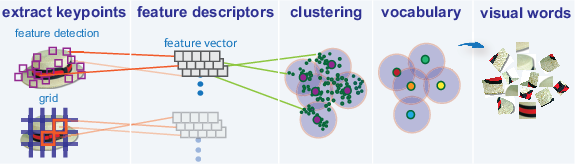
+ Cột ngang là các visual words.

+ Hàng dọc là số tần suất, trọng số của visual word



**Hình 3.1 Bag of visual words.**

## Các bước thực hiện:



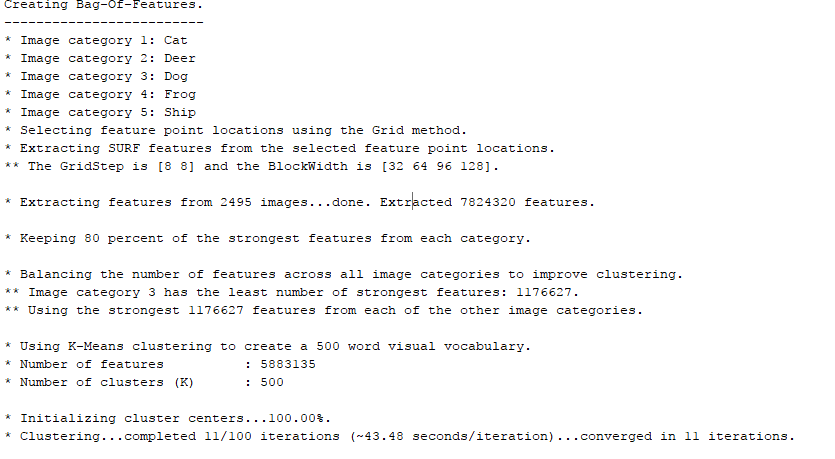
**Hình 3.2 Quá trình xây dựng bag of word**

Bước 1: Chuẩn bị dữ liệu train và test, sử dụng hàm imageDatastore để quản lý dữ liệu là gồm các ảnh.

Bước 2: Xây dựng Bag Of Word, ta dùng hàm bagOfFeatures trong matlab (matlab 2017a). Trong hàm bagOfFeatures thì sẽ gồm các xử lý:

- Trích xuất các đặc trưng SURF từ tập dữ liệu train.

- Dùng thuật giải gom cụm (clustering) như là K-Means trên tập hợp vector local feature thành các cụm riêng biệt, mỗi cụm là một feature hay visual word.



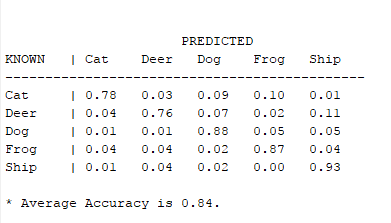
Hình 3.3 Xây dựng bag of feature trong matlab

Bước 3: Dùng trainImageCategoryClassifier để trả về một image classifier.

Bước 4: Hiển thị kết quả, dùng hàm evaluate trong matlab.

## Kết quả nhận xét:

Các tham số mặc định: VocabularySize = 500 ,GridStep = [8 8], BlockWidth – Patch size to extract SURFT descriptor = [32 64 96 128].

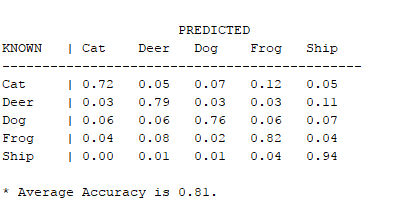


Hình 3.4 Bag Of Feature với các tham số mặc định

Áp dụng CustomExtractor: CustomExtractor = exampleBagOfFeaturesExtractor – là một hàm có sẵn trong MATLAB, thu được **8393180** feature.

*extractorFcn = @exampleBagOfFeaturesExtractor;*

*bag = bagOfFeatures(imgSets,'CustomExtractor',extractorFcn)*



Hình 3.5 Bag of Feature với CustomExtractor

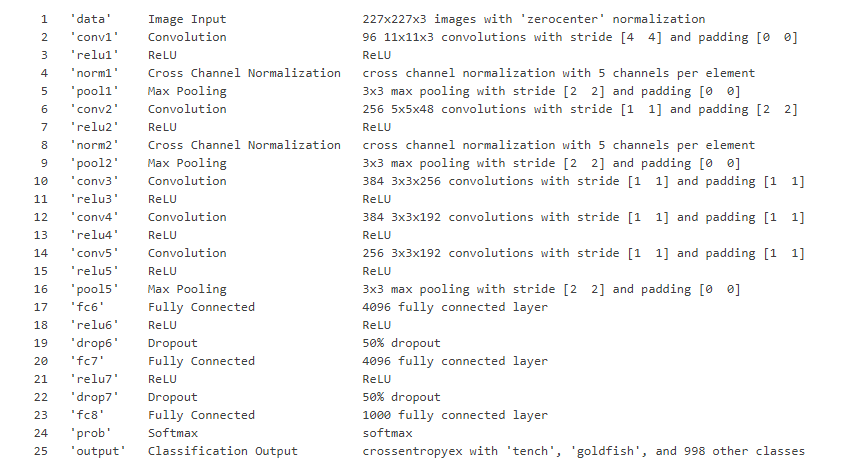
## Phương pháp 3:

Yêu cầu: Tải và cài đặt Neural Network Toolbox Model for AlexNet Network support package[8].

## Các bước thực hiện:

Bước 1: Tương tự, chuẩn bị dữ liệu train và test, sử dụng hàm imageDatastore để quản lý dữ liệu là gồm các ảnh.

Bước 2: Load Pre-trained AlexNet Network, các layer trong alexnet:



Ta chọn một số layer để kiểm thử như: ‘fc7’ – *Fully Connected, 4096 fully connected layer*, ‘fc8’ – *Fully connected, 1000 fully connected layer*, ‘conv4’- *Convolution: 384 3x3x192 convolutions with stride [1 1] and padding [1 1]*.

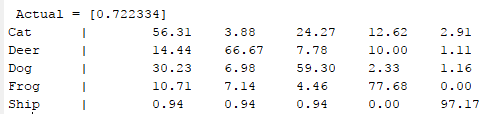
Thư viện alexnet chỉ có thể xử lý ảnh mà có input size = [ 227 227 3], nên ta dùng hàm *imresize* để xử lý.

Bước 3: Dùng hàm *fitcecoc* để trả về một classifier

Bước 4: Đánh giá độ chính xác và hiển thị kết quả phân loại trên bộ dữ liệu test

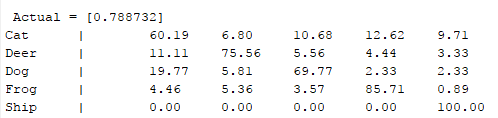
## Kết quả nhận xét:

Áp dụng tham số: feature layer= ‘conv4’



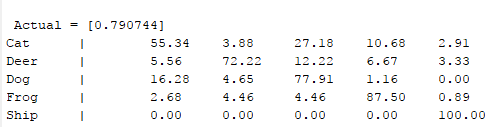
Hình 3.6 Áp dụng tham số feature layer = ‘conv4’

Áp dụng tham số: feature layer = ‘fc8’



Hình 3.7 Áp dụng feature layer = ‘fc8’

Áp dụng tham số feature layer = ‘fc7’



Hình 3.8 Áp dụng feature layer = ‘fc7’

## Đánh giá:

Bảng tóm tắt:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| parameter | [4 4] | [8 8] | fc7 | fc8 | conv4 | default value argument | custom feature extractor |
| Average Accuracy | 46.9 | 43.7 | 79.1 | 78.9 | 72.2 | 84.32 | 80.1 |

Qua kết quả thực nghiệm, ta thấy độ chính xác của phương pháp sử dụng HOG features thấp hơn so với hai phương pháp còn lại.

# KẾT LUẬN

Trong quá trình hoàn thành báo cáo, với thời gian và lượng kiến thức trong lĩnh vực này có hạn cùng với những phức tạp trong việc xử lý dữ liệu đầu vào, tuy nhiên báo cáo cũng đã giúp hiểu được quy trình phát triển hệ thống phân loại ảnh (image classification). Bên cạnh đó, em cũng hiểu thêm về công cụ matlab và xử lý ảnh.

# HƯỚNG PHÁT TRIỂN

- Phát triển giao diện thân thiện với người dùng.

- Thực nghiệm với nhiều tham số khác nhau hơn.

- Áp dụng thêm nhiều phương pháp khác nhau.

# DANH MỤC TÀI LIỆU THAM KHẢO

[1]- *Bài giảng môn Nhận dạng thị giác và ứng dụng,* TS. Lê Đình Duy

[2]- *Bài thực hành môn Nhận dạng thị giác và ứng dụng*, TS. Nguyễn Tấn Trần Minh Khang

[3]- *Image Processing Toolbox™ User's Guide* © COPYRIGHT 1993–2015 by The Math Works, Inc.

[4]- https://www.cs.toronto.edu/~kriz/cifar.html

[5]- http://www.mathworks.com

[6]- https://www.mathworks.com/help/vision/examples.html#d0e23

[7]-https://www.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html

[8]- https://www.mathworks.com/help/nnet/ref/alexnet.html

[9]-https://www.mathworks.com/help/vision/examples/image-category-classification-using-deep-learning.html

[*http://stackoverflow.com/*](http://stackoverflow.com/)