BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC CẦN THƠ

KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG

**🙞 🕮 🙜**



**BÁO CÁO ĐỒ ÁN**

**MÔN THỊ GIÁC MÁY TÍNH**

**Đề tài**

**NHẬN DẠNG HÌNH ẢNH**

**(Nhận dạng sử dụng đặc trưng cục bộ + mô hình Bag of Words**

**+ máy học libSVM)**

**Nhóm 06**

**Huỳnh Bảo Quốc**

**Lê Kim Yến**

**Nguyễn Trung Hậu**

**Nguyễn Thị Như Thảo**

**Khóa : 40**

Cần Thơ, 05/2018

BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC CẦN THƠ

KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG

🙞 🕮 🙜



**BÁO CÁO ĐỒ ÁN**

**MÔN THỊ GIÁC MÁY TÍNH**

**Đề tài**

**NHẬN DẠNG HÌNH ẢNH**

**(Nhận dạng sử dụng đặc trưng cục bộ +mô hình Bag of Words**

**+ máy học libSVM)**

**Giáo viên hướng dẫn: Nhóm Sinh viên thực hiện:**

**PGS.Ts Phạm Nguyên Khang Huỳnh Bảo Quốc B1400516**

**Lê Kim Yến B1400542**

**Nguyễn Trung Hậu B1400493**

**Nguyễn Thị Như Thảo B1400524**

**Khóa : 40**

Cần Thơ, 5/2018

**NHẬN XÉT CỦA GIẢNG VIÊN**

--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

**LỜI CẢM ƠN**

Lời đầu tiên chúng em xin cảm ơn thầy Phạm Nguyên Khang đã tận tình hướng dẫn môn Thị giác máy tính, để chúng em có được những kiến thức cũng như kĩ năng lập trình về python, ý tưởng giải quyết vấn đề trong đề tài.

Chúng em cũng xin cảm ơn bạn bè đã hỗ trợ giải quyết khó khăn trong quá trình làm bài đồ án này.

Chúng em chân thành cảm ơn thầy!

Cần Thơ, ngày 18 tháng 5 năm 2018

Nhóm sinh viên thực hiện

**MỤC LỤC**

[PHẦN GIỚI THIỆU 4](#_Toc514357787)

[**1. Đặt vấn đề: 4**](#_Toc514357788)

[**2. Lịch sử giải quyết vấn đề: 4**](#_Toc514357789)

[**3. Mục tiêu đề tài: 4**](#_Toc514357790)

[**4. Phạm vi nghiên cứu: 5**](#_Toc514357791)

[**5. Phương pháp nghiên cứu: 5**](#_Toc514357792)

[**6. Kết quả đạt được: 5**](#_Toc514357793)

[**7. Bố cục Đồ án 5**](#_Toc514357794)

[PHẦN NỘI DUNG 6](#_Toc514357795)

[CHƯƠNG 1 MÔ TẢ BÀI TOÁN 6](#_Toc514357796)

[**1. Mô tả chi tiết bài toán: 6**](#_Toc514357797)

[**2. Chức năng: 6**](#_Toc514357798)

[**3. Mô tả, giới thiệu công nghệ thực hiện: 6**](#_Toc514357799)

[**3.1. Phần mềm Pycharm: 6**](#_Toc514357800)

[**3.2. Ngôn ngữ Python: 7**](#_Toc514357801)

[**3.3. Sử dụng thư viên máy học libSVM: 7**](#_Toc514357802)

[**3.4. Sử dụng thư viện OpenCV: 7**](#_Toc514357803)

[CHƯƠNG 2 8](#_Toc514357804)

[THIẾT KẾ VÀ CÀI ĐẶT 8](#_Toc514357805)

[**1. Thiết kế: 8**](#_Toc514357806)

[**2. Cài đặt giải thuật: 8**](#_Toc514357807)

[**3. Giao diện hệ thống: 9**](#_Toc514357808)

[CHƯƠNG 3 12](#_Toc514357809)

[KIỂM THỬ VÀ ĐÁNH GIÁ 12](#_Toc514357810)

[**1. Mục tiêu: 12**](#_Toc514357811)

[**2. Kịch bản kiểm thử: 12**](#_Toc514357812)

[**3. Kết quả kiểm thử: 12**](#_Toc514357813)

[PHẦN KẾT LUẬN 13](#_Toc514357814)

[**1. Kết quả đạt được: 13**](#_Toc514357815)

[**2. Hướng phát triển:ss 13**](#_Toc514357816)

[TÀI LIỆU THAM KHẢO 14](#_Toc514357817)

**DANH MỤC HÌNH**

[Hình 1:Nhận dạng cây trồn 4](#_Toc514356418)

[Hình 2:Nhận dạng khuôn mặt trên facebook 4](#_Toc514356419)

[Hình 3: Phần mềm Pycharm IDE 7](#_Toc514356420)

[Hình 4: Python 7](#_Toc514356421)

[Hình 5: Giao diện của công cụ nhận diện ảnh 10](#_Toc514356422)

[Hình 6: Giao diện khi tải ảnh lên 11](#_Toc514356423)

**DANH MỤC BẢNG**

[Bảng 1: Mô tả giao diện 9](#_Toc514165290)

**TÓM TẮT**

Trong thời đại công nghệ hiện nay thì không thể không nhắc đến trí tuệ nhân tạo. Trí tuệ nhân tạo là một trong những lĩnh vực được các chuyên gia nghiên cứu để phát triển và sáng tạo ra các dự án như robot thông minh hay ngôi nhà thông minh,… và trí tuệ nhân tạo cũng được ứng dụng rộng rãi trong thị giác máy tính. Nói chung, Thị giác máy tính là một môn học khoa học liên quan đến lý thuyết đằng sau các hệ thống nhân tạo có trích xuất các thông tin từ các hình ảnh. Dữ liệu hình ảnh có thể nhiều dạng, chẳng hạn như chuỗi video, các cảnh từ đa camera, hay dữ liệu đa chiều từ máy quét y học. Thị giác máy tính còn là một môn học kỹ thuật, trong đó tìm kiếm việc áp dụng các mô hình và các lý thuyết cho việc xây dựng các hệ thống thị giác máy tính. Hơn nữa, Thị giác máy tính có rất nhiều giải thuật cũng như mô hình để áp dụng nó vào các công cụ hay là các ứng dụng. Trong đó có giải thuật SIFT là một giải thuật giúp chúng ta trích xuất đặc trưng trong mỗi ảnh, đồng thời kết hợp với mô hình Bag of Words để phân cụm cho các trích xuất cuối cùng là thư viện của máy học libSVM để có thể nhận diện hình ảnh hay đối tượng chính xác. Dùng phần mềm Pycharm và thư viện Opencv để cài đặt giải thuật SIFT cộng với mô hình BoW và thư viện libSVM trong máy học và được viết trên ngôn ngữ python để tạo nên công cụ nhận diện hình ảnh hay là những đối tượng xung quanh chúng ta.

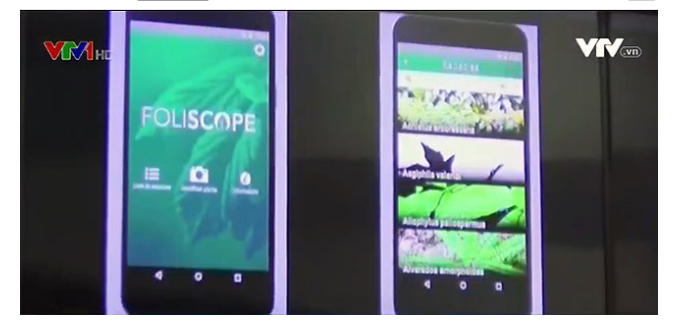
# PHẦN GIỚI THIỆU

## Đặt vấn đề:

Hiện nay, trí tuệ nhân tạo đang là xu hướng phát triển trong thời đại công nghệ 4.0 cũng không thể không nhắc đến một nhánh nhỏ trong đó chính là thị giác máy tính.Thị giác máy tính được áp dụng trong nhiều lĩnh vực như: sinh học, vật lý, khoa học máy tính,…. Xử lý ảnh, nhận diện ảnh là một trong những vấn đề mà được rất nhiều người quan tâm đến và nó giúp nhận diện các đối tượng trong đời sống một cách hiệu quả.Trong đồ án lần này, nhóm chúng em sẽ xây dựng công cụ nhận diện các đối tượng trong đời sống.

## Lịch sử giải quyết vấn đề:

* Vấn đề đã được giải quyết với nhiều cách khác nhau trên thế giới.
* Một vài ví dụ về ứng dụng như sau: mình như [Google Photos](https://tinhte.vn/tags/google-photos-2/), Google Search, [Google Drive](https://tinhte.vn/tags/google-drive/)… để tối ưu hóa việc phát hiện hình ảnh thông qua các từ khóa tìm kiếm của người dùng.



Hình 1:Nhận dạng cây trồng

Các nhà khoa học tại Costa Rica đang phát triển ứng dụng trên điện thoại, cho phép người sử dụng có thể nhận dạng cây trồng chỉ bằng cách chụp lại ảnh của lá cây đó.



Hình 2:Nhận dạng khuôn mặt trên facebook

1. **Mục tiêu đề tài:**

* Nhận diện đối tượng ảnh

1. **Phạm vi nghiên cứu:**

* Đối tượng của đề tài: nghiên cứu công cụ nhận diện hình ảnh cho phép mọi người đưa hình ảnh từ ngoài vào để công cụ nhận diện và đưa ra kết quả.
* Phạm vi nghiên cứu:
* Ứng dụng chạy trên máy tính.
* Các ứng dụng liên quan đến nhận diện ảnh.
* Tìm hiểu kỹ thuật lập trình trên Pycharm, ngôn ngữ Python,thư viện OpenCV.

## Phương pháp nghiên cứu:

* Phương pháp thu thập thông tin

+ Học hỏi thông tin từ bạn bè

+ Internet

* Về lý thuyết:

Sử dụng ngôn ngữ Python, kĩ thuật lập trình, các kiến thức đã học

* Về kỹ thuật:

+ Sử dụng ngôn ngữ Python

+ Sử dụng phầm mềm Pycharm IDE

## Kết quả đạt được:

* Số lượng lớp: 5 lớp được nhận diện
* Độ chính xác của mô hình khoảng 82,6%

## 7. Bố cục Đồ án

Giới thiệu tổng quát về đề tài: đặt vấn đề, lịch sử giải quyết vấn đề, mục tiêu đề tài, đối tượng phạm vi nghiên cứu, phương pháp nghiên cứu, kết quả đạt được, bố cục đồ án.

**Giới thiệu tổng quát về đề tài.**

Phần nội dung

Chương 1:Mô tả bài toán.

Chương 2: Thiết kế, cài đặt .

Chương 3: Kiểm thử và đánh giá bài toán.

Phần kết luận

Trình bày kết quả đạt được và hướng phát triển của giải thuật.

# PHẦN NỘI DUNG

# CHƯƠNG 1 MÔ TẢ BÀI TOÁN

## Mô tả chi tiết bài toán:

**Bước 1:** Thu thập ảnh từ Internet

**Bước 2:** Chuẩn bị dataset:

2.1: trích xuất đặc trưng tất cả các ảnh training

2.2: Sử dụng mô hình BoW để tiến hành phân các đặc trưng thành K cụm và lưu K cụm lại

2.3: Tiến hành trích xuất đặc trưng trên từng ảnh và tiến hành so sánh với cụm đã tìm được ở bước trên để xem từng feature thuộc cụm nào. Kết quả bước lại là 1 vector gồm K chiều, từng phần tử của nó là số lượng feature thuộc về cụm tương ứng

**Bước 3:** Training với libSVM

3.1: Chia tập dataset ra làm 2/3 training và 1/3 testing

3.2 : Tiến hành training với tập 2/3 train và dự đoán với 1/3 test, đo độ chính xác của mô hình và lưu mô hình

**Bước 4:** Test

4.1: Load lại mô hình ở bước 3

4.2 : Trích xuất đặc trưng trên ảnh cần dự đoán, ta cũng tiến hành so sánh với cụm ở bước 2.3 và cho ra 1 vector tương tự

4.3: Sử dụng mô hình của libSVM đã lưu và dự đoán kết quả

## Chức năng:

Nhận diện năm đối tượng như: máy bay, thuyền buồm, đồng hồ, cây kiểng, rùa biển.

## Mô tả, giới thiệu công nghệ thực hiện:

* 1. **Phần mềm Pycharm:**

Pycharm một công cụ dùng để lập trình python, pycharm có rất nhiều tiện dụng trong lập trình như là có các gợi ý trong lúc lập trình, đổi tên tập tin một cách nhanh chóng,….



Hình 3: Phần mềm Pycharm IDE

* 1. **Ngôn ngữ Python:**



Hình 4: Python

Ban đầu, Python được phát triển để chạy trên nền [Unix](https://vi.wikipedia.org/wiki/Unix). Nhưng rồi theo thời gian, nó đã "bành trướng" sang mọi [hệ điều hành](https://vi.wikipedia.org/wiki/H%E1%BB%87_%C4%91i%E1%BB%81u_h%C3%A0nh) từ [MS-DOS](https://vi.wikipedia.org/wiki/MS-DOS) đến [Mac OS](https://vi.wikipedia.org/wiki/Mac_OS), OS/2, [Windows](https://vi.wikipedia.org/wiki/Microsoft_Windows), [Linux](https://vi.wikipedia.org/wiki/Linux) và các hệ điều hành khác thuộc họ Unix. Mặc dù sự phát triển của Python có sự đóng góp của rất nhiều cá nhân, nhưng Guido van Rossum hiện nay vẫn là tác giả chủ yếu của Python.

* 1. **Sử dụng thư viên máy học libSVM:**

Đây là một thư viện đơn giản, dễ sử dụng và hiệu quả đối với phân lớp SVM và hồi qui. Nó có thể giải quyết phân lớp C-SVM, nu-SVM, hồi qui epsilon-SVM và hồi qui nu-SVM. Nó cũng cung cấp một công cụ lựa chọn mô hình tự động đối với phân lớp C-SVM. Tài liệu này giải thích cách sử dụng thư viện LIBSVM

* 1. **Sử dụng thư viện OpenCV:**

OpenCV (Thư viện hình ảnh máy tính nguồn mở) được phát hành theo giấy phép BSD và do đó nó hoàn toàn miễn phí cho cả mục đích học thuật và thương mại. Nó có giao diện C ++, Python và Java và hỗ trợ Windows, Linux, Mac OS, iOS và Android. OpenCV được thiết kế cho hiệu quả tính toán và tập trung mạnh vào các ứng dụng thời gian thực. Được viết bằng C / C ++ được tối ưu hóa, thư viện có thể tận dụng lợi thế của xử lý

# CHƯƠNG 2

# THIẾT KẾ VÀ CÀI ĐẶT

## Thiết kế:

Bài báo cáo bao gồm ba tập tin cài đặt:

* Main\_training.py: tập tin này sẽ chạy qua các qúa trình
  + Trích xuất tất cả các ảnh
  + Training BoW để ra các cụm(Centroids)
  + Tính các vector histogram cho từng ảnh thông qua các Centroids trên
  + Training với libSVM và lưu mô hình
* Main\_nhandang.py tập tin này sẽ chạy qua các quá trình
  + Trích xuất ảnh cần dự đoán
  + Tính vector thông qua các Centroids
  + Dùng mô hình SVM đã lưu để dự đoán vector trên
* Utils.py: tập tin này sẽ soạn tất cả các hàm thông dụng tái sử dụng nhiều lần như : đọc file, lấy danh sách tên tập tin, tính toán các đặc trưng thuộc cụm và cho ra một vector tương ứng

1. **Cài đặt giải thuật:**

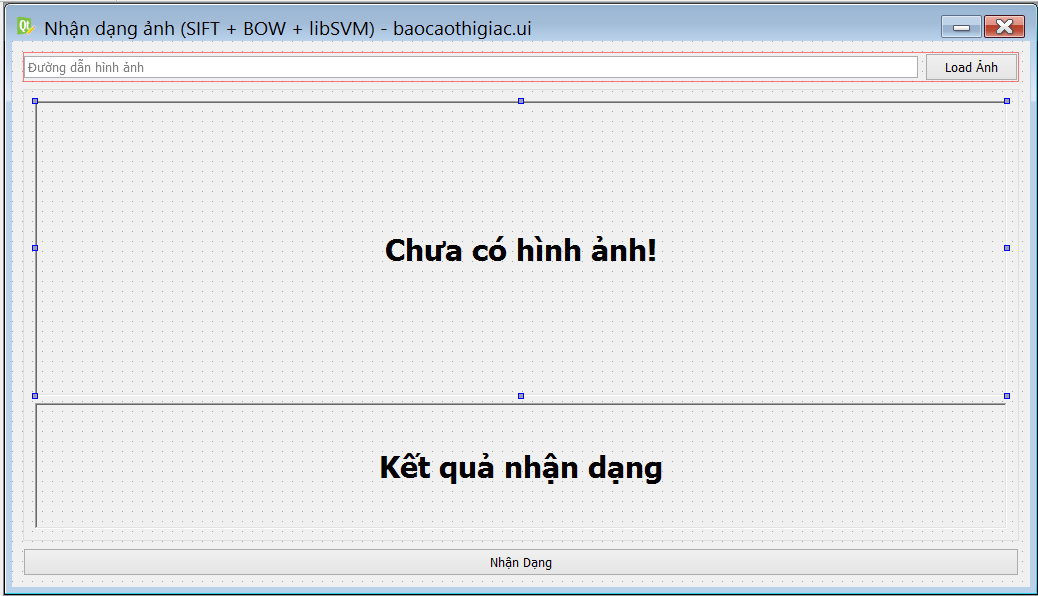
* Cài dặt SIFT: SIFT là phát minh đã được đăng kí bản quyền của David Lowe.
* Mỗi véc tơ mô tả là một ma trận 4x4 các tổ chức đồ. Mỗi tổ chức đồ có 8 khoảng tương ứng với 8 hướng. Do đó, mỗi véc tơ mô tả SIFT là một véc tơ 4x4x8=128 chiều. Lúc này, mỗi ảnh được biểu diễn bởi một tập các véc tơ mô tả SIFT.
* Cài đặt BoW:
  + Mô hình bag-of-words là một biểu diễn đơn giản được sử dụng trong xử lý ngôn ngữ tự nhiên và truy xuất thông tin (IR). Còn được gọi là mô hình không gian vectơ . Trong mô hình này, một văn bản (chẳng hạn như một câu hoặc một tài liệu) được đại diện như là túi của các từ của nó, bỏ qua ngữ pháp và thậm chí cả thứ tự từ nhưng giữ đa dạng. Mô hình túi-của-chữ cũng đã được sử dụng cho thị giác máy tính. Mô hình bag-of-words thường được sử dụng trong các phương pháp phân loại tài liệu, trong đó tần suất xuất hiện của mỗi từ được sử dụng như một tính năng để huấn luyện một trình phân loại,
* Sử dụng libSVM
* Máy học khá phức tạp với nhiều bài viết về chúng, nhưng hầu hết là khó hiểu. Nhóm muốn viết cho dễ hiểu và minh hoạ cách dùng libSVM và ứng dụng của nó.
* Nó là thư viện miễn phí. Về căn bản, nó cho phép bạn lấy vài mẫu dữ liệu đã có, huấn luyện để tạo 1 mô hình nhận dạng, và rồi sử dụng mô hình này để dự đoán.

## Giao diện hệ thống:

* Mô tả sơ lược về giao diện

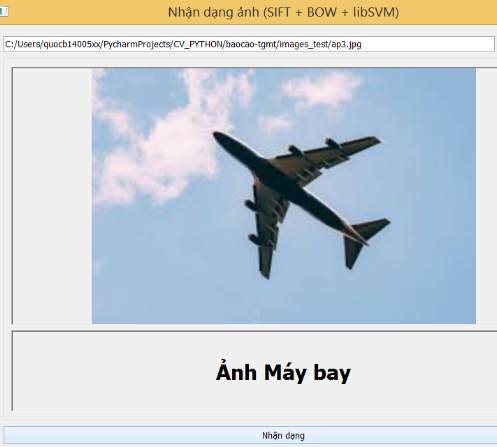
|  |  |  |
| --- | --- | --- |
| Tên | Thanh điều khiển | Công dụng |
| Load ảnh |  | Tải ảnh từ máy tính lên công cụ |
| Nhận dạng |  | Khi tải ảnh lên bắm phím nhận dạng trả về kết quả xem ảnh đó là gì. |

Bảng 1: Mô tả giao diện



Hình 5: Giao diện của công cụ nhận diện ảnh

* + - Nếu như chưa tải ảnh lên sẽ hiện thị chưa có ảnh , và khi có ảnh thì sẽ có kết quả trả về đó là tên của ảnh



Hình 6: Giao diện khi tải ảnh lên

* + - Khi lấy một tấm ảnh bất kì, và nhận dạng thì kết quả trả về là tên của đối tượng đã tải lên.

# CHƯƠNG 3

# KIỂM THỬ VÀ ĐÁNH GIÁ

## Mục tiêu:

* Tải ảnh lên máy sẽ trả kết quả ảnh tên ảnh đó .

## Kịch bản kiểm thử:

* Kiểm tra 10 ảnh trên mỗi đối tượng, để kiểm tra sự đúng đắn của mỗi ảnh với các lớp

## Kết quả kiểm thử:

* Với 5 đối tượng, kiểm tra 50 ảnh,hiệu quả lên đến từ 45 cho đến 48 ảnh

# 

# PHẦN KẾT LUẬN

## 1. Kết quả đạt được:

* Thu lại một kết quả đúng khi qua quá trình kiểm thử
* Độ chính xác khá cao đó là 82,6% với 5 lớp mỗi lớp 500 ảnh (2500 thể hiện)

## 2. Hướng phát triển:

* Tăng cường lớp dữ liệu.
* Đảm báo chất lượng mô hình với độ chính xác cao.
* Trang bị máy tính có cấu hình tốt để quá trình huấn luyện dữ liệu tiết kiệm được nhiều thời gian.

# TÀI LIỆU THAM KHẢO

1. PGS.Ts Phạm Nguyên Khang, *Silde bài giảng môn Thị giác máy tính.* Địa chỉ

[***https://drive.google.com/drive/folders/14iDbxMoWnYkAYHVCygm5ieurzUnrcqX***](https://drive.google.com/drive/folders/14iDbxMoWnYkAYHVCygm5ieurzUnrcqX)

1. Wikipedia tiếng Việt, *Bách khoa toàn thư mở tiếng Việt* Địa chỉ: <https://vi.wikipedia.org/wiki/Th%E1%BB%8B_gi%C3%A1c_m%C3%A1y_t%C3%ADnh>
2. Tìm hiểu về SIFT và OpenCV. Địa chỉ

[*http://opencvpythontutroals.readthedocs.io/en/latest/py\_tutorials/py\_feature2d/py\_sift\_intro/py\_sift\_intro.html*](http://opencvpythontutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html)