**ĐẠI HỌC SƯ PHẠM KỸ THUẬT THÀNH PHỐ HỒ CHÍ MINH**

**KHOA ĐÀO TẠO CHẤT LƯỢNG CAO**

**

**Đồ Án 3**

**Đề tài:**

**Nhận Diện Điểm Số**

|  |
| --- |
| **SVTH: NGUYỄN VĂN HÀ** |
| **MSSV: 17110130** |
| **SVTH: NGUYỄN VĂN THẮNG** |
| **MSSV: 17110230** |
| **Khoá: 2017** |
| **Ngành: CÔNG NGHỆ THÔNG TIN** |
| **GVHD:**  **TS. Trần Tiến Đức** |

Thành phố Hồ Chí Minh, tháng 01 năm 2021

**LỜI CẢM ƠN**

Lời đầu tiên, nhóm thực hiện xin được phép gửi lời cảm ơn chân thành đến khoa Đào tạo Chất Lượng Cao – Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh đã tạo mọi điều kiện thuận lợi nhất cho nhóm thực hiện có cơ hội được tự do tiếp cận, tham khảo, mở rộng thêm kiến thức trong lĩnh vực Công nghệ thông tin nói chung và trong quá trình thực hiện đồ án 3 nói riêng!

Lời cảm ơn trân trọng nhất nhóm thực hiện xin chân thành gửi đến thầy Trần Tiến Đức đã cùng đồng hành và trực tiếp giảng dạy, hướng dẫn và tạo mọi điều kiện thuận lợi giúp đỡ cho nhóm phát huy hết khả năng cũng như nâng cao kiến thức trong suốt quá trình học tập, đặc biệt là trong quá trình chuẩn bị và thực hiện môn học Học Máy. Cảm ơn sự nhiệt tình của Cô, là động lực vô cùng to lớn giúp nhóm thực hiện, kiên trì trong suốt quá trình thực hiện đề tài và khám phá ra những kiến thức mới đầy thú vị và bổ ích liên quan đến python machine learning mà cụ để là đề tài Nhận Dạng Điểm Số.

**Nhóm sinh viên thực hiện**

**Nguyễn Văn Hà**

**Nguyễn Văn Thắng**

**LỜI MỞ ĐẦU**

Ngày nay công nghệ thông tin đã và đang phát triển đạt đến mức độ cao, được ứng dụng vào mọi ngành, mọi lĩnh vực của nền kinh tế đất nước góp phần quan trọng làm cho các ngành này phát triển nhanh hơn, hiệu quả hơn. Việc tìm hiểu về machine learning trong thời đại công nghệ 4.0 này nhằm đáp ứng nhu cầu ngày càng cao của người dùng trên nhiều lĩnh vực.

Bên cạnh nhu cầu về trí tuệ nhân tạo nhận dạng các vật thể càng được năng cao và phát triển hơn trong thời đại 4.0.Việc nhập điểm thi vào cơ sở dữ liệu là một điều có thể thay thế được để tiết kiệm thời gian và công sức chính vì vậy tụi em đã nghiên cứu tìm hiểu và chọn đề tài Nhận diện điểm số từ file ảnh

Mục Lục

[LỜI CẢM ƠN i](#_Toc64112942)

[LỜI MỞ ĐẦU ii](#_Toc64112943)

[Chương 1: TỔNG QUAN 1](#_Toc64112944)

[1.1 Lý do chọn đề tài 1](#_Toc64112945)

[1.2. Mục tiêu nghiên cứu đề tài 1](#_Toc64112948)

[Chương 2: NỘI DUNG CHÍNH 2](#_Toc64112949)

[2.1. Bài toán object detection 2](#_Toc64112952)

[2.2. Thuật Toán CNN 3](#_Toc64112953)

[2.2.1. Khái Niệm 3](#_Toc64112954)

[2.2.2. Cấu trúc của mạng CNN 5](#_Toc64112955)

[Cách chọn tham số cho CNN 6](#_Toc64112956)

[Chương 3: Demo Thuật Toán 8](#_Toc64112957)

[3.1.1. Opencv 9](#_Toc64112959)

[Chương 4: TỔNG KẾT 11](#_Toc64112960)

[4.1. Kết luận 11](#_Toc64112964)

[4.2. Ưu điểm 11](#_Toc64112965)

[4.3. Nhược điểm 11](#_Toc64112966)

[TÀI LIỆU THAM KHẢO 11](#_Toc64112967)

# Chương 1: TỔNG QUAN

## Lý do chọn đề tài

Machine learning là một lĩnh vực của [trí tuệ nhân tạo](https://vi.wikipedia.org/wiki/Tr%C3%AD_tu%E1%BB%87_nh%C3%A2n_t%E1%BA%A1o) liên quan đến việc nghiên cứu và xây dựng các kĩ thuật cho phép các hệ thống "học" tự động từ dữ liệu để giải quyết những vấn đề cụ thể. Ví dụ như các máy có thể "học" cách phân loại [thư điện tử](https://vi.wikipedia.org/wiki/Th%C6%B0_%C4%91i%E1%BB%87n_t%E1%BB%AD) xem có phải [thư rác (spam)](https://vi.wikipedia.org/wiki/Th%C6%B0_r%C3%A1c) hay không và tự động xếp thư vào thư mục tương ứng. Học máy rất gần với [suy diễn thống kê](https://vi.wikipedia.org/wiki/Suy_di%E1%BB%85n_th%E1%BB%91ng_k%C3%AA) .

Ngoài ra việc xử lý ảnh bằng python sau đó tiến hành việc phân tích dữ liệu, nhưng khác với thống kê, học máy tập trung vào sự phức tạp của các giải thuật trong việc thực thi tính toán.Vì thế một phần của học máy là nghiên cứu sự phát triển các giải thuật suy luận xấp xỉ mà có thể xử lý được . Chính vì thế với kiến thức học tụi em đã nghiên cứu và chọn đề tài tìm hiểu và thực hiện nhận dạng điểm số.



## Mục tiêu nghiên cứu đề tài

Tìm hiểu và sử dụng thuật toán để nhận dạng xử lý ảnh bằng python

**Chương 2: NỘI DUNG CHÍNH**







## Bài toán object detection

a/ Object dection là gì?

* Hiểu một cách nôm na, object detection là một bài toán:

+ Phân loại hình ảnh (image classification): liên quan đến việc gán nhãn cho hình ảnh.

+ Định vị vật thể (object localization): liên quan đến việc vẽ một hộp giới hạn (bounding box) xung quanh một hoặc nhiều đối tượng trong hình ảnh nhằm khoanh vùng đối tượng.

* + Phát hiện đối tượng (object detection): Là nhiệm vụ khó khăn hơn và là sự kết hợp của cả hai nhiệm vụ trên: Vẽ một bounding box xung quanh từng đối tượng quan tâm trong ảnh và gán cho chúng một nhãn. Kết hợp cùng nhau, tất cả các vấn đề này được gọi là object recognition hoặc object detection.

b/ Như thế nào là nhận dạng đối tượng?

* Nhận dạng đối tượng là một thuật ngữ chung để mô tả một tập hợp các nhiệm vụ thị giác máy tính có liên quan liên quan đến việc xác định các đối tượng trong ảnh kỹ thuật số.
* Phân loại hình ảnh liên quan đến việc dự đoán lớp của một đối tượng trong một hình ảnh. Định vị vật thể đề cập đến việc xác định vị trí của một hoặc nhiều đối tượng trong một hình ảnh và vẽ bounding box xung quanh chúng. Phát hiện đối tượng kết hợp hai nhiệm vụ trên và thực hiện cho một hoặc nhiều đối tượng trong hình ảnh. Chúng ta có thể phân biệt giữa ba nhiệm vụ thị giác máy tính cơ bản trên thông qua input và output của chúng như sau:

+ Phân loại hình ảnh: Dự đoán nhãn của một đối tượng trong một hình ảnh.

* + Input: Một hình ảnh với một đối tượng, chẳng hạn như một bức ảnh.
  + Output: Nhãn lớp (ví dụ: một hoặc nhiều số nguyên được ánh xạ tới nhãn lớp).

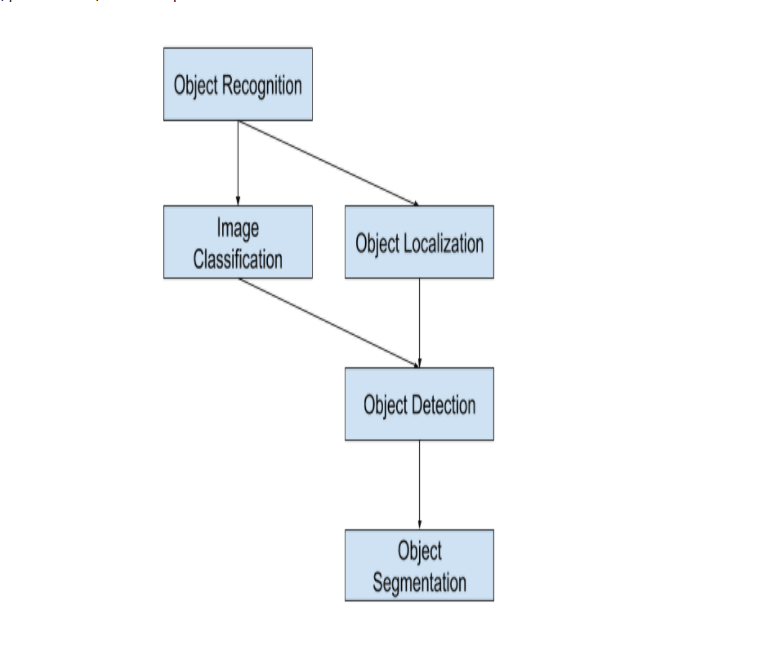
+ Định vị đối tượng: Xác định vị trí hiện diện của các đối tượng trong ảnh và cho biết vị trí của chúng bằng bounding box.

* + Input: Một hình ảnh có một hoặc nhiều đối tượng, chẳng hạn như một bức ảnh.
  + Output: Một hoặc nhiều bounding box được xác định bởi tọa độ tâm, chiều rộng và chiều cao.

+ Phát hiện đối tượng: Xác định vị trí hiện diện của các đối tượng trong bounding box và nhãn của các đối tượng nằm trong một hình ảnh.

* + Input: Một hình ảnh có một hoặc nhiều đối tượng, chẳng hạn như một bức ảnh.
  + Output: Một hoặc nhiều bounding box và nhãn cho mỗi bounding box.

Một số định nghĩa khác cũng rất quan trọng trong computer vision là phân đoạn đối tượng (object segmentation), trong đó các đối tượng được nhận dạng bằng cách làm nổi bật các pixel cụ thể của đối tượng thay vì bounding box. Và image captioning kết hợp giữa các kiến trúc mạng CNN và LSTM để đưa ra các lý giải về hành động hoặc nội dung của một bức ảnh.



## Thuật Toán CNN

### Khái Niệm

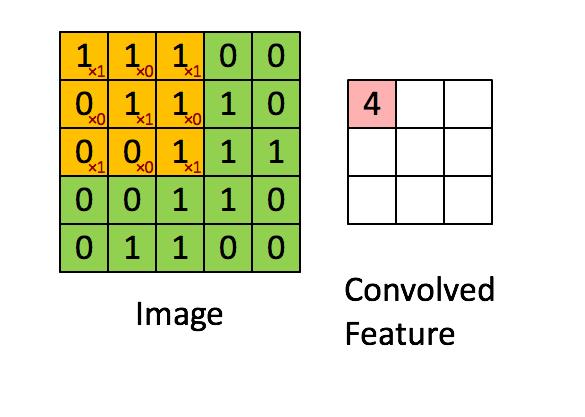
**A) Convolutional Neural Network là gì?**

Convolutional Neural Network (CNNs – Mạng nơ-ron tích chập) là một trong những mô hình Deep Learning tiên tiến. Nó giúp cho chúng ta xây dựng được những hệ thống thông minh với độ chính xác cao như hiện nay.

CNN được sử dụng nhiều trong các bài toán nhận dạng các object trong ảnh. Vậy nên tại sao thuật toán này được sử dụng rộng rãi cho việc nhận dạng (detection), chúng ta hãy cùng tìm hiểu về thuật toán này.

**B) Convolutional là gì?**

Là một cửa sổ trượt (Sliding Windows) trên một ma trận như mô tả hình dưới:



Convolutional layer có các parameter(kernel) đã được học để tự điều chỉnh lấy ra những thông tin chính xác nhất mà không cần chọn các feature.

Trong hình ảnh ví dụ trên, ma trận bên trái là một hình ảnh trắng đen được số hóa. Ma trận có kích thước 5×5 và mỗi điểm ảnh có giá trị 1 hoặc 0 là giao điểm của dòng và cột.

Convolution hay tích chập là nhân từng phần tử trong ma trận 3. Sliding Window hay còn gọi là kernel, filter hoặc feature detect là một ma trận có kích thước nhỏ như trong ví dụ trên là 3×3.

### Cấu trúc của mạng CNN

Mạng CNN là một tập hợp các lớp Convolution chồng lên nhau và sử dụng các hàm nonlinear activation như ReLU và tanh để kích hoạt các trọng số trong các node. Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo.

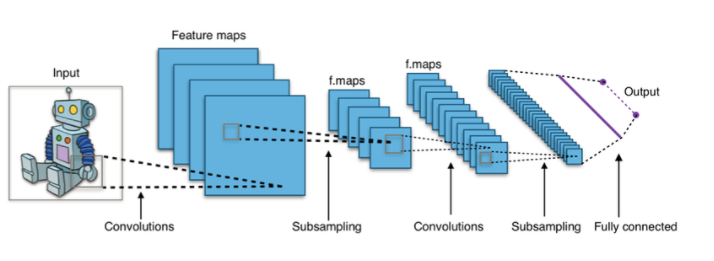
Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo. Trong mô hình mạng truyền ngược (feedforward neural network) thì mỗi neural đầu vào (input node) cho mỗi neural đầu ra trong các lớp tiếp theo.

Mô hình này gọi là mạng kết nối đầy đủ (fully connected layer) hay mạng toàn vẹn (affine layer). Còn trong mô hình CNNs thì ngược lại. Các layer liên kết được với nhau thông qua cơ chế convolution.

Layer tiếp theo là kết quả convolution từ layer trước đó, nhờ vậy mà ta có được các kết nối cục bộ. Như vậy mỗi neuron ở lớp kế tiếp sinh ra từ kết quả của filter áp đặt lên một vùng ảnh cục bộ của neuron trước đó.

Mỗi một lớp được sử dụng các filter khác nhau thông thường có hàng trăm hàng nghìn filter như vậy và kết hợp kết quả của chúng lại. Ngoài ra có một số layer khác như pooling/subsampling layer dùng để chắt lọc lại các thông tin hữu ích hơn (loại bỏ các thông tin nhiễu).

Trong quá trình huấn luyện mạng (traning) CNN tự động học các giá trị qua các lớp filter dựa vào cách thức mà bạn thực hiện. Ví dụ trong tác vụ phân lớp ảnh, CNNs sẽ cố gắng tìm ra thông số tối ưu cho các filter tương ứng theo thứ tự raw pixel > edges > shapes > facial > high-level features. Layer cuối cùng được dùng để phân lớp ảnh.



Trong mô hình CNN có 2 khía cạnh cần quan tâm là **tính bất biến** (Location Invariance) và **tính kết hợp** (Compositionality). Với cùng một đối tượng, nếu đối tượng này được chiếu theo các gốc độ khác nhau (translation, rotation, scaling) thì độ chính xác của thuật toán sẽ bị ảnh hưởng đáng kể.

Pooling layer sẽ cho bạn tính bất biến đối với phép dịch chuyển (translation), phép quay (rotation) và phép co giãn (scaling). Tính kết hợp cục bộ cho ta các cấp độ biểu diễn thông tin từ mức độ thấp đến mức độ cao và trừu tượng hơn thông qua convolution từ các filter.

Đó là lý do tại sao CNNs cho ra mô hình với độ chính xác rất cao. Cũng giống như cách con người nhận biết các vật thể trong tự nhiên.

1. **Mạng CNN sử dụng 3 ý tưởng cơ bản:**

* **các trường tiếp nhận cục bộ** (local receptive field)
* **trọng số chia sẻ** (shared weights)
* **tổng hợp** (pooling).

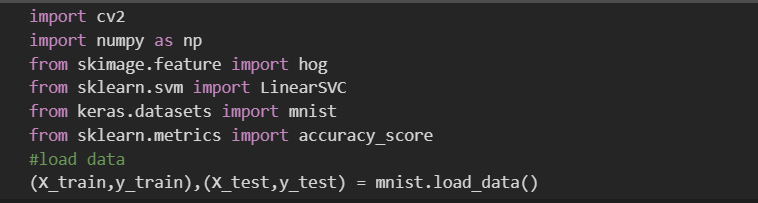
## Cách chọn tham số cho CNN

1. Số các convolution layer: càng nhiều các convolution layer thì performance càng được cải thiện. Sau khoảng 3 hoặc 4 layer, các tác động được giảm một cách đáng kể
2. Filter size: thường filter theo size 5×5 hoặc 3×3
3. Pooling size: thường là 2×2 hoặc 4×4 cho ảnh đầu vào lớn
4. Cách cuối cùng là thực hiện nhiều lần việc train test để chọn ra được param tốt nhất.

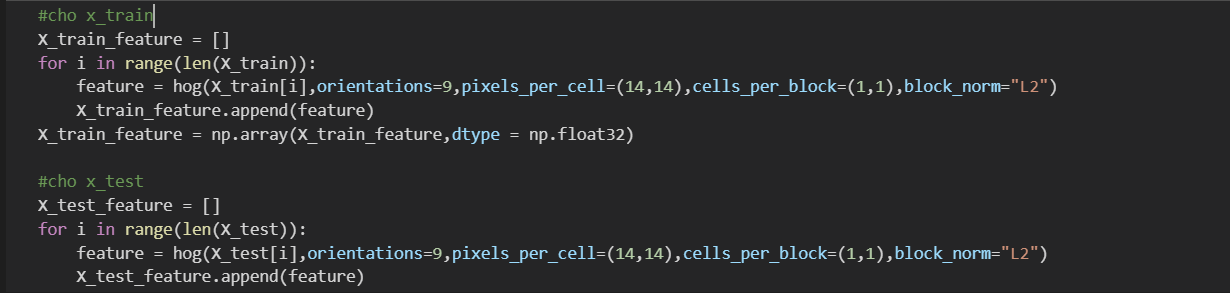
# Chương 3: Demo Thuật Toán

3.1. Một số bước tiêu biểu

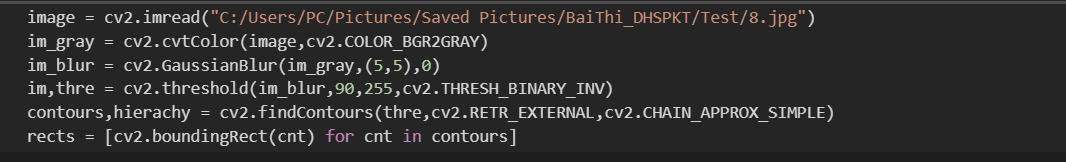
* Khai báo các thư viện cần thiết và load data vào



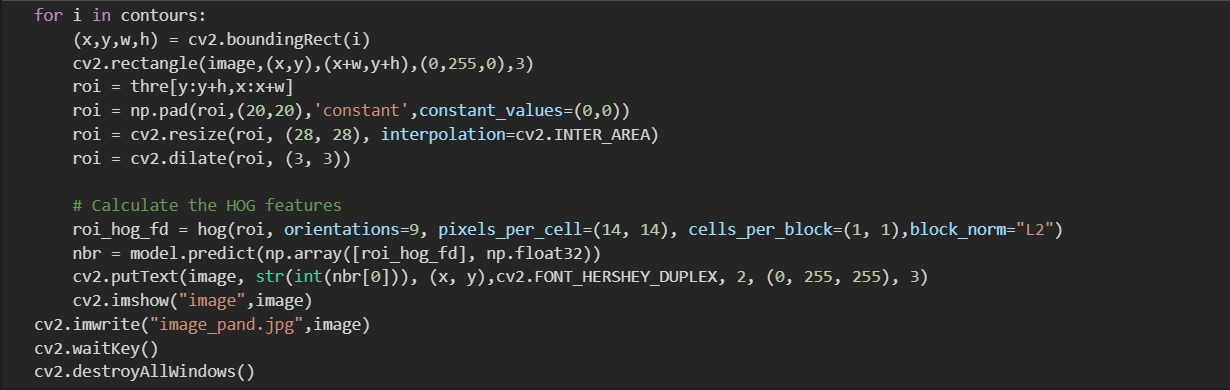
Chia data thành 2 phần rồi đem đi train và test



* Nạp hình ảnh và xử lý



* Tính toán vector HOG





### Opencv

OpenCV là một thư viện mã nguồn mở hàng đầu cho thị giác máy tính (computer vision), xử lý ảnh và máy học, và các tính năng tăng tốc GPU trong hoạt động thời gian thực.

Python: Ngôn ngữ được dùng nhiều để demo / test OpenCV do tính ngắn gọn, ít phải thiết lập. Bên cạnh đó, nếu dùng Python thì cũng có thể code được trên nhiều hệ điều hành.

* 1. **Các ứng dụng OpenCV**

OpenCV đang được sử dụng rộng rãi trong các ứng dụng bao gồm:

Hình ảnh street view Kiểm tra và giám sát tự động Robot và xe hơi tự lái Phân tích hình ảnh y tế Tìm kiếm và phục hồi hình ảnh/video Phim - cấu trúc 3D từ chuyển động Nghệ thuật sắp đặt tương tác

**Chức năng OpenCV**

Image/video I/O, xử lý, hiển thị (core, imgproc, highgui) Phát hiện các vật thể (objdetect, features2d, nonfree) Geometry-based monocular or stereo computer vision (calib3d, stitching, videostab) Computational photography (photo, video, superres) Machine learning & clustering (ml, flann) CUDA acceleration (gpu)

**2. Cài đặt thư viện.**

Chúng ta sẽ làm việc thông qua nhiều ví dụ Python ở đây. Bắt đầu với ràng buộc Python của OpenCV . Bạn sẽ cần hai thư viện chính, với một tùy chọn thứ ba: python-OpenCV, Numpy, và Matplotlib. Chúng ta bắt đầu cài đặt trên Linux nhé.

apt-get install python3-pip

sau đó chạy

pip3 install numpy or apt-get install python3-numpy

Tiếp theo chạy lệnh

pip3 install matplotlib or apt-get install python3-matplotlib

Create 1 file với đuôi py nhé. import các thư viện cần thiết vào

import cv2

import matplotlib

# Chương 4: TỔNG KẾT







## Kết luận

Đã tìm hiểu và áp dụng những kiến thức về bài toán nhận dạng điểm số

## Ưu điểm

* Chương trình sử dụng các thư viện hổ trợ mạnh mẽ như hog,opencv,…
* Đề tài thực tế phù hợp trong tình tràng hiện nay

## Nhược điểm

* Chương trình vẫn còn chưa hoàn thiện.
* Một vài chỗ xử lý logic chưa được chính xác.
* Giao diện chưa được bắt mắt.
* Vẫn chưa đọc được với tỉ lệ chính xác cao

**TÀI LIỆU THAM KHẢO**



[1] Applied Supervised Learning with Python

[2] <https://machinelearningcoban.com/2016/12/28/linearregression/#-gioi-thieu>

[3] https://viblo.asia/p/nhan-dang-anh-co-ban-voi-python-bWrZn6mbZxw