



Object Recognition

*E-mail: pima.vn@gmail.com

Mô tả dự án

Image Recognition là một trong những bài toán khá quan trọng và điển hình trong xử lý ảnh. **Object Recognition** cùng với **image classification** và **image segmentation** là nền tảng cho ngành Thị giác Máy tính (**computer vision**) và mang nhiều ý nghĩa rất quan trọng. Mục tiêu của bài toán này là **xác định** (các) đối tượng trong ảnh, trong đó việc xác định được định nghĩa như sau:

- (1) Xác định vị trí đối tượng
- (2) Xác định lớp (class) của đối tượng

Là một bài toán kinh điển và được đặt ra câu hỏi từ rất sớm, thế nhưng mạng học sâu **convolutional neural network (CNN)** với các kiến trúc biến thể đã chứng minh được sự mạnh mẽ so với các phương thức truyền thống. Ở dự án này, các bạn sẽ nghiên cứu các kiến trúc mạng tích chập để giải quyết bài toán Object Recognition.

Câu hỏi gợi ý

- (1) Mô tả chi tiết nội dung của các lớp cơ bản (convolution, pooling, batch normalize..) trong kiến trúc mạng cơ bản của CNN.
- (2) Hãy trình bày mô hình CNN thích hợp để giải quyết bài toán Image Recognition. Hàm mục tiêu cần tối ưu trong mô hình này là gì?
- (3) Nghiên cứu cách hoạt động của Region CNN (RCNN) và các biến thể của nó (Fast RCNN, Faster RCNN) trong bài toán Image Recognition.
- (4) Giải thích khái niệm về region proposal, Region of Interest (ROI) pooling, Anchor... và các khái niệm mới trong RCNN so với CNN nguyên bản.
- (5) Thực nghiệm, quan sát và nhận xét kết quả trên tập dữ liệu Stanford Dogs.
 - (a) Initialization có ảnh hưởng thế nào đến kết quả cuối cùng?
 - (b) Số lớp có ảnh hưởng thế nào đến quá trình huấn luyện? Số tham số tự do thì sao?
 - (c) Thứ tự huấn luyện ưu tiên như thế nào là hợp lý?
 - (d) Data preprocessing có giúp ích gì không?

Kiến thức: Ma trận, các phép toán cơ bản trên ma trận, không gian \mathbb{R}^n , vi tích phân nhiều biến, bài toán tìm cực trị, học sâu, mạng thần kinh tích chập

Một số từ khóa: Recognition, Detection, Classification, Intersection over Union, Non-maxima suppression, Hyperplane, Optimization, Feature detector, Feature map, Weight sharing, RCNN

Tham Khảo

-
- [1] Các bài giảng PiMA 2019
 - [2] <https://www.kaggle.com/jessicali9530/stanford-dogs-dataset>