



# Image Segmentation

\*E-mail: [pima.vn@gmail.com](mailto:pima.vn@gmail.com)

## Mô tả dự án

**Image Segmentation** là một trong những bài toán khá quan trọng và điển hình trong xử lý ảnh. **Image Segmentation** cùng với **object recognition** và **image classification** là nền tảng cho ngành **computer vision** và mang nhiều ý nghĩa rất quan trọng. Mục tiêu của bài toán này là **phân chia bức ảnh ra thành các vùng**, trong đó các điểm ảnh thuộc mỗi vùng mang cùng một màu sắc tương ứng với nhãn dán của đối tượng đó, các vùng kề nhau mang màu sắc tương phản khá rõ rệt, tìm được đúng hình dạng của các đối tượng. Trước những năm 2000, các phương pháp xử lý cổ điển thông dụng có thể kể đến là lọc ngưỡng, tìm cạnh, gom cụm, v.v. Đến năm 2010, **convolutional neural network (CNN)** với khả năng học các tính năng phân cấp, đã khẳng định mình là công cụ mạnh mẽ và hiệu quả vượt trội 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 Image Segmentation.

## 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 UNet trong bài toán Image Segmentation.
- (4) Giải thích khái niệm về encoder và decoder, phép transposed convolution,... và các khái niệm mới so với CNN nguyên bản.
- (5) Thực nghiệm, quan sát và nhận xét kết quả trên tập dữ liệu trong Carvana Image Masking Challenge.
  - (a) Initialization có ảnh hưởng thế nào đến kết quả cuối cùng
  - (b) Số lớp có ảnh hưởng thế nào đến quá trình train? Số tham số tự do thì sao?
  - (c) 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, kiến trúc mạng thần kinh tích chập

**Một số từ khóa:** Instance Segmentation, Semantic Segmentation, Optimization, Feature detector, Translation invariance, Weight sharing, Encoder, Decoder, Unet

## Tham Khảo

- [1] Các bài giảng PiMA 2019
- [2] <https://www.kaggle.com/c/carvana-image-masking-challenge>