

# Phân đoạn đối tượng ngụy trang bằng các mô hình học sâu

Ngô Phúc Danh

Trường Đại học Công nghệ Thông tin,  
ĐHQG-HCM

## Tóm tắt

Nghiên cứu tập trung vào việc làm nổi bật các đặc điểm của đối tượng ngụy trang, từ đó phân đoạn chi tiết từng vùng ảnh của các đối tượng.

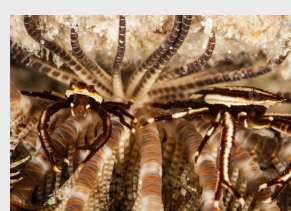
- One-stage (OSFormer): thực hiện phân đoạn và dự đoán đối tượng ngụy trang trong một lượt duy nhất.
- Two-stage (DCNet): bao gồm hai giai đoạn với giai đoạn đầu để xác định đối tượng ngụy trang và giai đoạn hai thực hiện phân đoạn chi tiết.

## Động lực

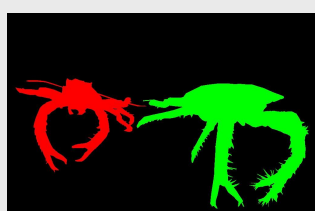
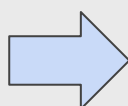
Giải quyết bài toán Camouflaged Instance Segmentation (CIS) chính là giải quyết các thách thức của con người đối với các vấn đề như:

- Quan sát tập tính, hành vi của các loài động vật có cơ chế ngụy trang trong môi trường tự nhiên.
- Phát hiện các loài động vật nguy hiểm ẩn mình.
- Phát hiện kẻ thù đang ngụy trang trên chiến trường.

## Tổng quan

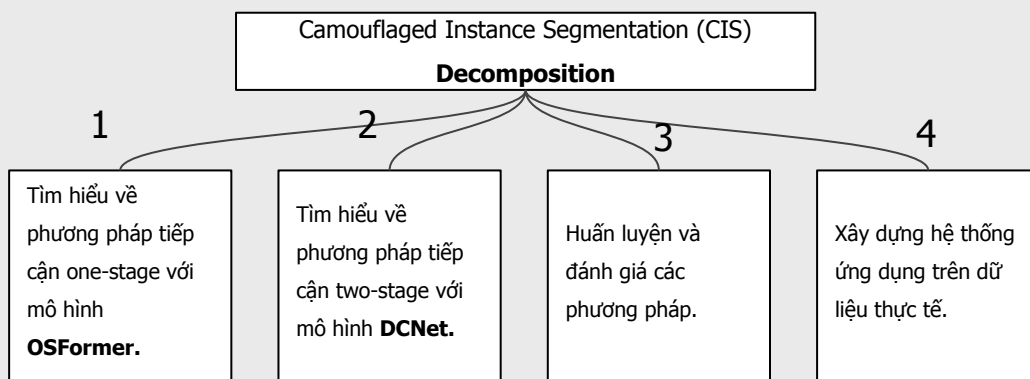


Input: ảnh chứa đối tượng ngụy trang.



Output: segmentation mask mô tả vị trí và hình dạng.

Hình 1. Minh họa input và output của bài toán CIS.



Hình 2. Sơ đồ minh họa tổng quan phương pháp.

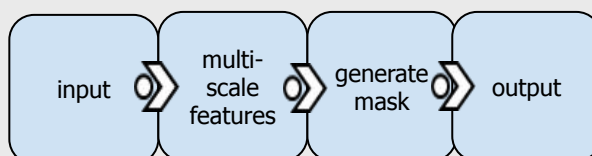
## Mô tả

### 1. Nội dung nghiên cứu

- Tìm hiểu về phương pháp tiếp cận one-stage với mô hình OSFormer.
- Tìm hiểu về phương pháp tiếp cận two-stage với mô hình DCNet.
- Tiến hành huấn luyện và đánh giá hai mô hình dựa trên các tiêu chí về độ chính xác và tốc độ inference.
- Xây dựng hệ thống ứng dụng trên dữ liệu thực tế.

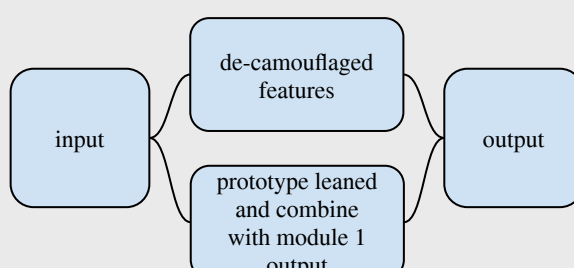
### 2. Phương pháp nghiên cứu

- One-stage (OSFormer): tìm hiểu từng components của kiến trúc có nhiệm vụ xử lý đặc trưng, kết hợp các thông tin high và low levels để tái tạo segmentation mask. Tìm hiểu các multi-scale features được kết hợp với nhau thông qua một quá trình encoder-decoder sử dụng cross-attention. Tìm hiểu về quá trình giám sát các đặc trưng cạnh trong quá trình training.



Hình 3. Flow nghiên cứu về phương pháp one-stage.

- Two-stage (DCNet): Tìm hiểu module đầu tiên đảm nhận nhiệm vụ khử đối tượng ngụy trang ở nhiều tỉ lệ khác nhau và module còn lại đảm nhận nhiệm vụ tái tạo segmentation mask bằng cách kết hợp với output của module đầu tiên. Tìm hiểu cách các multi-scale features khác nhau kết hợp với nhau thông qua kiến trúc FPN. Tìm hiểu các prototype được học để kết hợp với output của module đầu tiên nhằm tăng cường khả năng tái tạo segmentation mask.



Hình 4. Flow nghiên cứu về phương pháp two-stage.

- Cài đặt và tiến hành huấn luyện trên bộ dữ liệu COD10K.
- Tìm hiểu phương pháp và độ đo để đánh giá sự hiệu quả của các hướng tiếp cận trên phương diện về độ chính xác cũng như tốc độ inference.
- Tìm hiểu cách deploy một hệ thống có thể áp dụng vào dữ liệu thực tế.

### 3. Kết quả mong đợi

- One-stage (OSFormer): giúp nâng cao sự chính xác trong việc phát hiện đối tượng ngụy trang, đồng thời đáp ứng được tốc độ inference tốt.
- Two-stage (DCNet): tăng khả năng tái tạo segmentation mask, điều này mang lại sự chính xác cao nhưng sẽ đánh đổi lại tốc độ inference chậm hơn do cần xử lý hai quá trình.
- Xây dựng được hệ thống phân đoạn các đối tượng ngụy trang thực tế trong môi trường tự nhiên.