



THUYẾT MINH ĐỀ CƯƠNG NGHIÊN CỨU

NGHIÊN CỨU PHƯƠNG PHÁP PHÂN ĐOẠN ẢNH TRONG ĐIỀU KIỆN ÍT DỮ LIỆU HUẤN LUYỆN

Ứng viên: Nguyễn Thành Danh

Cán bộ hướng dẫn: TS. Nguyễn Vinh Tiệp

PGS.TS. Nguyễn Văn Tâm

Nội dung

1. Giới thiệu ứng viên
2. Giới thiệu đề tài
3. Tổng quan tình hình nghiên cứu
4. Nội dung nghiên cứu đề xuất
5. Các kết quả sơ khởi
6. Kế hoạch thực hiện đề tài

1. Giới thiệu ứng viên

- **Quá trình học tập:**

- **Đại học:** Cử nhân Tài năng ngành Khoa học máy tính, Trường Đại học Công nghệ Thông tin (ĐH CNTT), ĐHQG TP.HCM
- **Cao học:** Thạc sĩ ngành Khoa học máy tính, Trường ĐH CNTT, ĐHQG TP.HCM

- **Đơn vị công tác:**

- Phòng Thí nghiệm Truyền thông Đa phương tiện, Trường ĐH CNTT, ĐHQG TP. HCM

- **Hướng nghiên cứu:**

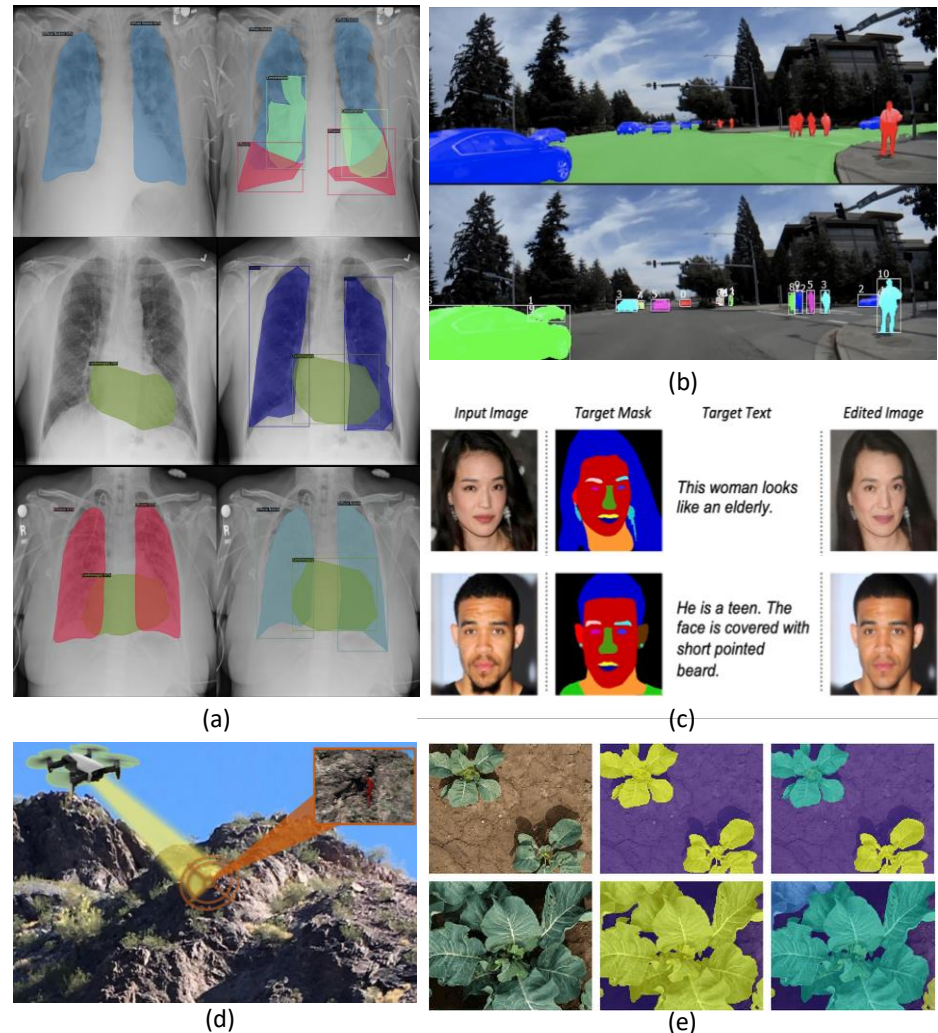
- Thị giác máy tính
- Các mô hình học sâu và ứng dụng

2. Giới thiệu đề tài

- **Phân đoạn ảnh:** phân tích nội dung hình ảnh với **mức độ chi tiết cao ở cấp độ điểm ảnh**
- **Phân đoạn ảnh có nhiều ứng dụng** trong thực tiễn: y khoa, giao thông thông minh...
- **Vấn đề:**
 - Chi phí gán nhãn cao
 - Thiếu dữ liệu gán nhãn, nhất là dữ liệu đặc thù
 - ...

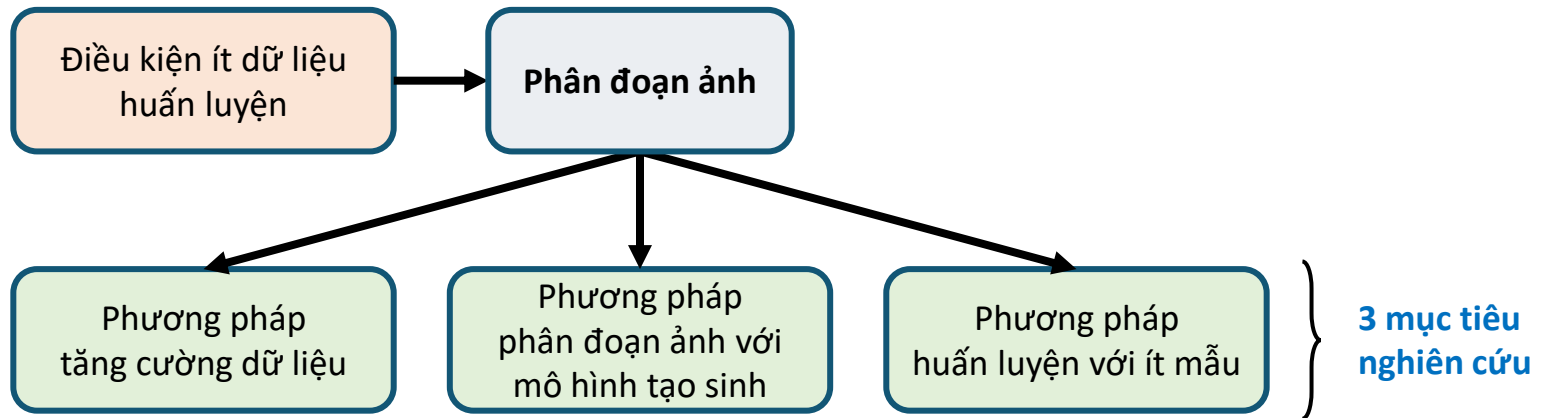
➔ Phân đoạn ảnh sử dụng ít dữ liệu gán nhãn?

Minh họa một số ứng dụng của bài toán phân đoạn ảnh



2. Giới thiệu đề tài

- Bài toán: **phân đoạn ảnh** trong **điều kiện ít dữ liệu huấn luyện**



- Các thách thức:**

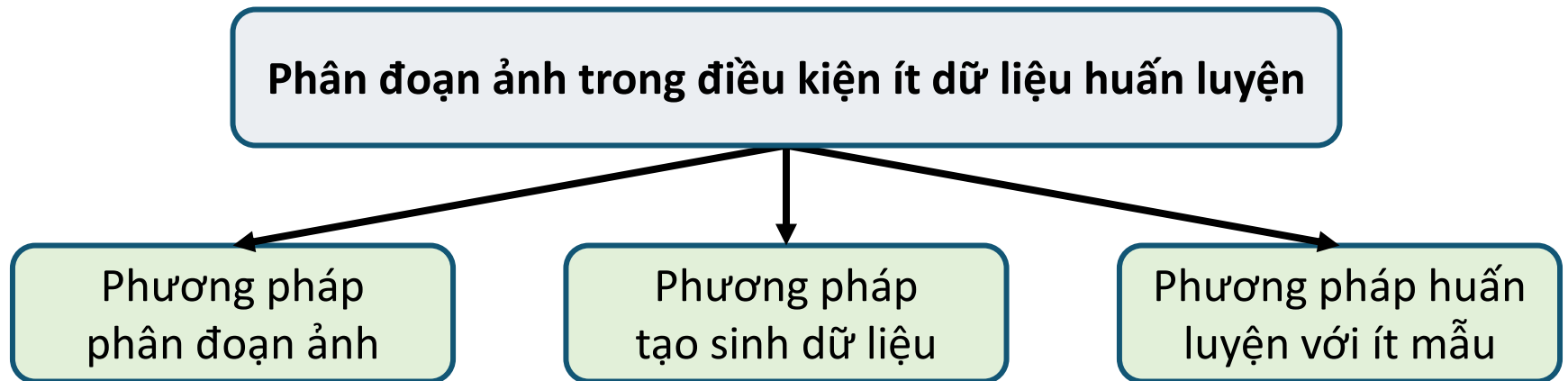
- Sự đa dạng của dữ liệu ảnh
- Hạn chế của mô hình CNN phân biệt truyền thống
- Chi phí gán nhãn dữ liệu

2. Giới thiệu đề tài – Mục tiêu nghiên cứu

Giải quyết bài toán **Phân đoạn ảnh** trong điều kiện ít dữ liệu huấn luyện

1. Nghiên cứu phương pháp **tăng cường dữ liệu** phục vụ cho bài toán **phân đoạn ảnh**
2. Nghiên cứu phương pháp **phân đoạn ảnh** với hướng tiếp cận dựa trên **mô hình tạo sinh dữ liệu**
3. Nghiên cứu phương pháp khai thác các mẫu dữ liệu gán nhãn với hướng tiếp cận **sử dụng ít mẫu huấn luyện**

3. Tổng quan tình hình nghiên cứu



3. Tổng quan tình hình nghiên cứu

Phương pháp phân đoạn ảnh

- **Hướng tiếp cận dựa trên cơ chế phân biệt^[1]**: mô hình thực hiện gán nhãn ngữ nghĩa cho từng điểm ảnh
 - Các mô hình 2 giai đoạn: họ Mask RCNN^[2], PANET^[3], HTC^[4]...
 - Các mô hình 1 giai đoạn: YOLACT^[5], OSFormer^[6], SAM^[7]...
 - **Hướng tiếp cận dựa trên cơ chế tạo sinh^[1]**: mô hình tạo sinh bản đồ phân đoạn ngữ nghĩa/ sử dụng thông tin từ mô hình tạo sinh
 - Các mô hình dựa trên GANs^[8]
 - Các mô hình dựa trên cơ chế khuếch tán Diffusion^{[9][10]}
- ➔ Khai thác được các thông tin từ **mô hình tạo sinh**
- ➔ Giải quyết hạn chế của hướng tiếp cận phân biệt: inductive bias

3. Tổng quan tình hình nghiên cứu

Các hướng tiếp cận tạo sinh dữ liệu có điều kiện:

- Dựa trên **dữ liệu đầu vào**:
 - Câu mô tả
 - Hình ảnh tham chiếu
 - Sử dụng meta data
 - Sử dụng hỗn hợp các điều kiện
- Dựa trên **kiến trúc mô hình**:
 - Cơ chế GANs^[11]
 - Cơ chế Diffusion^[12]
 - Cơ chế Progressive^[13]

➔ Phương pháp **tăng cường sự đa dạng cho dữ liệu**

[11] X. Pan, M. Zhang, D. Ding and M. Yang, "A Geometrical Perspective on Image Style Transfer With Adversarial Learning," in IEEE TPAMI, 2022.

[12] Kavar, Bahjat, Shiran Zada, Oran Lang, Omer Tov et al. "Imagic: Text-based real image editing with diffusion models." arXiv preprint, 2022.

[13] Yang, Shuai, Liming Jiang, Ziwei Liu, and Chen Change Loy. "Pastiche master: exemplarbased high-resolution portrait style transfer." In CVPR, 2022.

3. Tổng quan tình hình nghiên cứu

- **Phương pháp huấn luyện với ít mẫu:** sử dụng tri thức thuộc miền tổng quát để thực hiện tác vụ trên miền cụ thể
- **Cơ chế:** dựa trên các kiến trúc phục vụ bài toán học có giám sát, xây dựng pipeline cho ngữ cảnh học ít dữ liệu huấn luyện
- **Một số phương pháp tiêu biểu:** Meta R-CNN^[14], CRNet^[15], iFS R-CNN^[16]

➔ Phương pháp **giải quyết vấn đề đặc thù ít dữ liệu huấn luyện**

[14] Xiaopeng Yan, Ziliang Chen, Anni Xu, Xiaoxi Wang et al. “Meta R-CNN: Towards general solver for instance-level low-shot learning”. In ICCV. 2019.

[15] Weide Liu, Chi Zhang, Guosheng Lin, and Fayao Liu. “CRNet: Cross-Reference Networks for Few-Shot Segmentation”. In CVPR. June 2020.

[16] K. Nguyen and Sinisa Todorovic. “iFS-RCNN: An Incremental Few-shot Instance Segmenter”. In CVPR. 2022.

4. Nội dung nghiên cứu đề xuất

Nội dung 1: Giải quyết vấn đề thiếu dữ liệu gán nhãn phân đoạn

Nội dung 2: Phân đoạn ảnh dựa trên mô hình tạo sinh

Nội dung 3: Phân đoạn ảnh sử dụng ít dữ liệu huấn luyện

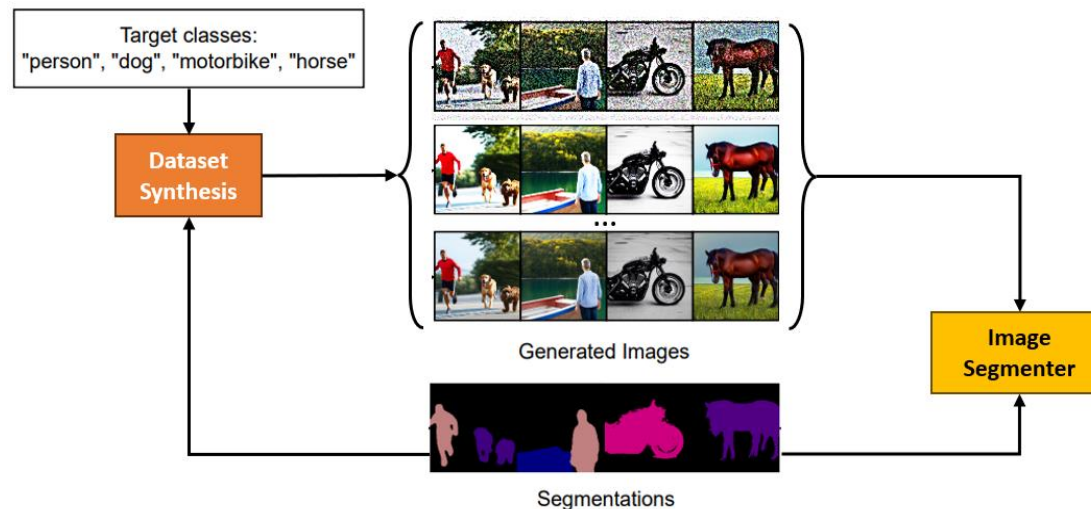
ND1: Giải quyết vấn đề thiếu dữ liệu gán nhãn phân đoạn

- **Động lực nghiên cứu:**

- Vấn đề thiếu dữ liệu gán nhãn cho bài toán phân đoạn ảnh
- Sự phát triển của các mô hình tạo sinh dữ liệu

- **Mục tiêu:**

- Sử dụng mô hình tạo sinh ảnh như một công cụ giúp tăng cường dữ liệu^[17]
- Khai thác hiệu quả các dữ liệu gán nhãn sẵn có



Minh họa ý tưởng tăng cường dữ liệu gán nhãn sử dụng mô hình tạo sinh dữ liệu có điều kiện

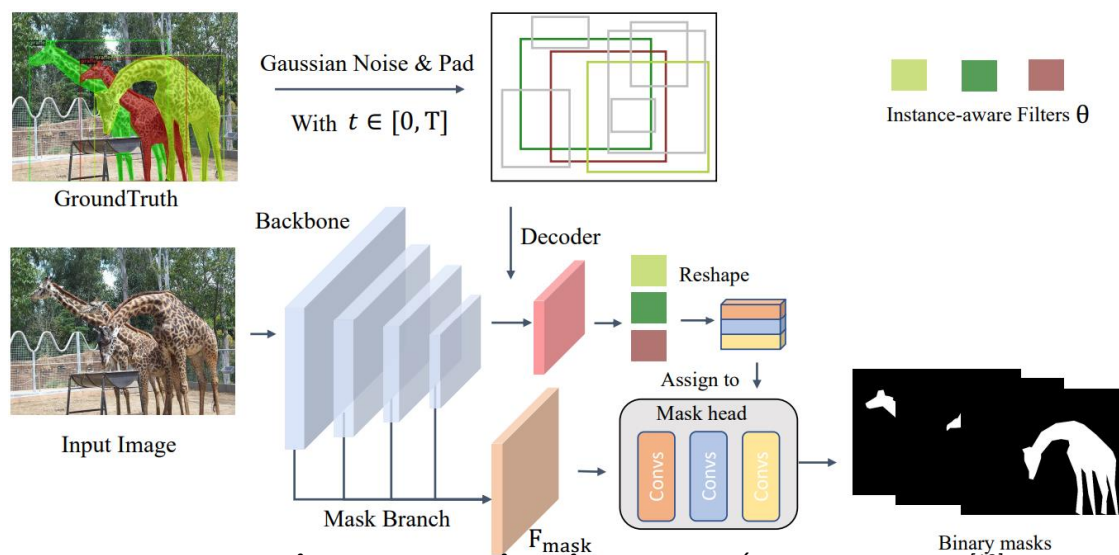
[17] Nguyen, Q., Vu, T., Tran, A., & Nguyen, K., Dataset diffusion: Diffusion-based synthetic data generation for pixel-level semantic segmentation. In NeurIPS 36. 2024

ND2: Phân đoạn ảnh dựa trên mô hình tạo sinh

- **Động lực nghiên cứu:**

- Khai thác thông tin giàu ngữ nghĩa từ các mô hình tạo sinh dữ liệu
- Vấn đề inductive bias của các mô hình phân biệt

- **Mục tiêu:** đề xuất mô hình áp dụng cơ chế tạo sinh dữ liệu để tăng cường thông tin phục vụ quá trình phân đoạn ảnh



Mình họa ý tưởng phân đoạn ảnh sử dụng cơ chế mô hình tạo sinh^[18]

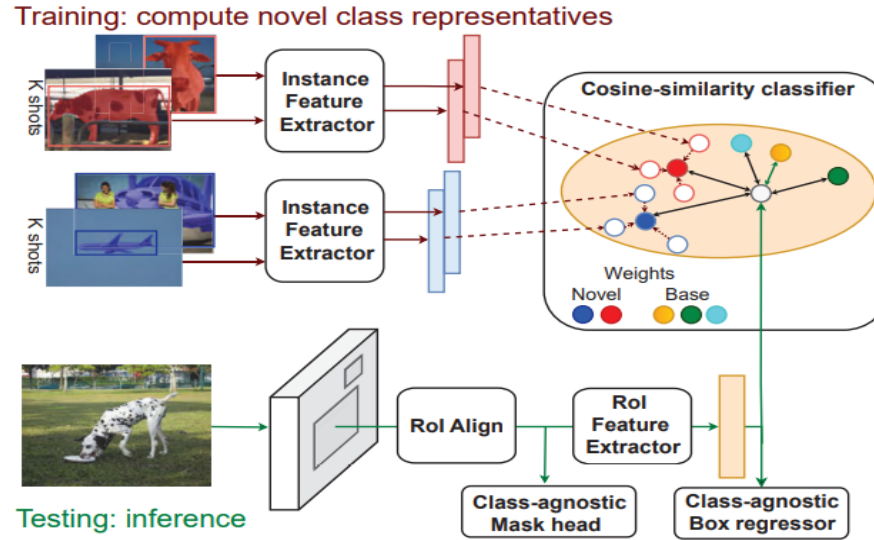
[18] Gu, Z., Chen, H., Xu, Z., Lan, J., Meng, C., & Wang, W. (2022). Diffusioninst: Diffusion model for instance segmentation. arXiv preprint arXiv:2212.02773.

ND3: Phân đoạn ảnh sử dụng ít dữ liệu huấn luyện

- **Động lực nghiên cứu:**

- Dựa trên quan sát thực tế về quá trình học hỏi tri thức của con người
- Khai thác hiệu quả số lượng ít mẫu dữ liệu gán nhãn

- **Mục tiêu:** Đề xuất mô hình khai thác phương pháp học với ít mẫu kết hợp sử dụng dữ liệu tạo sinh



Minh họa ý tưởng phân đoạn ảnh sử dụng phương pháp huấn luyện với ít mẫu dữ liệu

[18] Dan Andrei Ganea, Bas Boom, and Ronald Poppe. "Incremental Few-Shot Instance Segmentation". In CVPR. 2021, pp. 1185–1

5. Các kết quả sơ khởi

- Luận văn đại học và cao học:
 - Phân đoạn ảnh với phương pháp tương thích miền dữ liệu
 - Phân đoạn ảnh khai thác đặc trưng có tính phân biệt cao
- Các đề tài khoa học công nghệ (KHCN):
 - Chủ nhiệm 2 đề tài KHCN cấp cơ sở
 - Tham gia 2 đề tài cấp cơ sở và cấp ĐHQG khác
- Các công bố khoa học:
 - 01 bài tạp chí ISI Q1 [CT2]
 - 05 bài báo hội nghị quốc tế [CT1, CT3, CT4, CT5, CT6]

6. Kế hoạch thực hiện đề tài

Nội dung thực hiện	Thời gian thực hiện
Khảo sát, nghiên cứu một số phương pháp phân đoạn ảnh, tạo sinh ảnh có điều kiện. Xác định nội dung nghiên cứu chi tiết.	6 tháng
Nghiên cứu, đề xuất cải tiến phương pháp tạo sinh ảnh có điều kiện phục vụ bài toán phân đoạn ảnh.	6 tháng
Đề xuất cải tiến mô hình phân đoạn ảnh dựa trên các điểm yếu từ những mô hình đã khảo sát, tập trung vào hướng tiếp cận sử dụng cơ chế tạo sinh dữ liệu và sử dụng ít dữ liệu huấn luyện.	12 tháng
Tiến hành cài đặt, thực nghiệm và đánh giá các mô hình đề xuất. Tổng hợp, công bố các kết quả nghiên cứu tại các hội nghị, tạp chí thuộc danh mục theo quy định.	6 tháng
Viết báo cáo và chuẩn bị bảo vệ luận án	6 tháng

Một số công trình đã công bố

- [CT1] **Thanh-Danh Nguyen**, Duc-Tuan Luu, Vinh-Tiep Nguyen, and Thanh Duc Ngo, “CE-OST: Contour Emphasis for One-Stage Transformer-based Camouflage Instance Segmentation.” 2023 International Conference on Multimedia Analysis and Pattern Recognition (MAPR). IEEE, 2023.
- [CT2] Sy-Tuyen Ho, Manh-Khanh Huu Ngo, **Thanh-Danh Nguyen**, Nguyen Phan, Vinh-Tiep Nguyen, Thanh Duc Ngo, Duy-Dinh Le, Tam V. Nguyen, "Abstraction-Perception Preserving Cartoon Face Synthesis", Multimedia Tools and Applications (MTA), 2023. IF = 2.395 (SCI).
- [CT3] **Thanh-Danh Nguyen***, Anh-Khoa Nguyen Vu*, Nhat-Duy Nguyen*, Vinh-Tiep Nguyen, Thanh Duc Ngo, Thanh-Toan Do, Minh-Triet Tran, and Tam V. Nguyen, “Few-shot Camouflaged Animal Detection and Segmentation”, CV4Animals Workshop, CVPR 2022. (invited poster)
- [CT4] Manh-Khanh Ngo Huu, Vinh Quang Ngo, **Thanh-Danh Nguyen**, Vinh-Tiep Nguyen, and Thanh Duc Ngo. “Antique Photo Restoration and Colorization via Generative Model.” 2022 International Conference on Multimedia Analysis and Pattern Recognition (MAPR). IEEE, 2022. (Scopus)
- [CT5] Anh-Khoa Nguyen Vu, **Thanh-Danh Nguyen**, Vinh-Tiep Nguyen, and Thanh Duc Ngo, “DF-FSOD: A Novel Approach for Few-shot Object Detection via Distinguished Features.” 2021 International Conference on Multimedia Analysis and Pattern Recognition (MAPR). IEEE, 2021. (Scopus)
- [CT6] Ngoc-Khanh Nguyen, **Thanh-Danh Nguyen**, and Vinh-Tiep Nguyen. “A More Focus on Multi-degradation Method for Single Image Super-Resolution.” 2021 International Conference on Multimedia Analysis and Pattern Recognition (MAPR). IEEE, 2021. (Scopus)

Tài liệu tham khảo

- [1] Chen, J., Lu, J., Zhu, X., & Zhang, L. Generative semantic segmentation. In CVPR. 2023.
- [2] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. “Mask R-CNN”. In ICCV. 2017
- [3] Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia. “Path Aggregation Network for Instance Segmentation”. In CVPR. 2018.
- [4] Kai Chen, Jiangmiao Pang, Jiaqi Wang, Yu Xiong, Xiaoxiao Li, Shuyang Sun, Wansen Feng, Ziwei Liu, Jianping Shi, Wanli Ouyang, et al. “Hybrid task cascade for instance segmentation”. In CVPR. 2019
- [5] Daniel Bolya, Chong Zhou, Fanyi Xiao, and Yong Jae Lee. “YOLACT: Real-time Instance Segmentation”. In ICCV. 2019.
- [6] Jialun Pei, Tianyang Cheng, Deng-Ping Fan, He Tang, Chuanbo Chen, and Luc Van Gool. “OSFormer: One-Stage Camouflaged Instance Segmentation with Transformers”. In ECCV. 2022.
- [7] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R.. “Segment anything”. In ICCV. 2023 (pp. 4015-4026). [8] Xun, S., Li, D., Zhu, H., Chen, M., Wang, J., Li, J., ... & Huang, P. “Generative adversarial networks in medical image segmentation: A review”. Computers in Biology and Medicine 2022.
- [9] Gu, Z., Chen, H., Xu, Z., Lan, J., Meng, C., & Wang, W. (2022). Diffusioninst: Diffusion model for instance segmentation. arXiv preprint arXiv:2212.02773
- [10] De Ridder, V., Dey, B., Halder, S., & Van Waeyenberge, B. “SEMI-DiffusionInst: A Diffusion Model Based Approach for Semiconductor Defect Classification and Segmentation”. In 2023 International Symposium ELMAR



THUYẾT MINH ĐỀ CƯƠNG NGHIÊN CỨU

NGHIÊN CỨU PHƯƠNG PHÁP PHÂN ĐOẠN ẢNH TRONG ĐIỀU KIỆN ÍT DỮ LIỆU HUẤN LUYỆN

Ứng viên: Nguyễn Thành Danh

Cán bộ hướng dẫn: TS. Nguyễn Vinh Tiệp

PGS.TS. Nguyễn Văn Tâm