

NGHIÊN CỨU TỔNG QUÁT HÓA TỪ NGUỒN DỮ LIỆU ĐƠN MIỀN CHO BÀI TOÁN ĐẾM NGƯỜI TRONG ẢNH BẰNG PHƯƠNG PHÁP MPCOUNT

Môn học: Phương pháp nghiên cứu khoa học

Lớp: CS2205.SEP2025

GV: PGS.TS. Lê Đình Duy

HV: Nguyễn Thùy Linh – 250104012

Tóm tắt

- Lớp: : CS2205.SEP2025
- Link Github:
<https://github.com/lilalinhlan/ReasearchMethodology>
- Link YouTube video:
- Họ và tên: Nguyễn Thùy Linh
- MSHV: 250104012

Giới thiệu (1/2)

1. Bối cảnh nghiên cứu

- Đếm đám đông ứng dụng trong giám sát an ninh, tổ chức sự kiện, quản lý giao thông và quy hoạch đô thị thông minh
- Phương pháp chủ đạo hiện nay: Hồi quy bản đồ mật độ (Density map regression) dựa trên học sâu
- Ưu điểm: Đạt độ chính xác cao khi dữ liệu huấn luyện và kiểm tra có phân phối tương đồng



Hình 1. Ví dụ về ứng dụng Density map regression

Giới thiệu (2/2)

2. Vấn đề nghiên cứu

- Hiện tượng Dịch chuyển miền dữ liệu (Domain Shift): Hiệu suất giảm mạnh khi triển khai ở môi trường có góc quay, thời tiết hoặc ánh sáng khác biệt so với lúc huấn luyện
- Hạn chế của Domain Adaptation: Đòi hỏi dữ liệu từ miền đích, khó đáp ứng trong thực tế
- Thử thách SDG: Vấn đề label ambiguity giữa người và nền trong các bối cảnh khác nhau



Hình 2. Môi trường thực tế khác biệt với lúc huấn luyện

Mục tiêu

- Tổng hợp, phân tích các kỹ thuật Domain Generalization hiện có và đánh giá hạn chế của chúng trong crowd counting
- Triển khai khung mô hình MPCount tích hợp cơ chế AMB và PC để trích xuất đặc trưng bất biến và giảm sai số nhãn.
- Đánh giá định lượng trên các bộ dữ liệu tiêu chuẩn (SHA, SHB, JHU-Crowd++) để chứng minh khả năng tổng quát hóa

Nội dung và Phương pháp (1/3)

1. Nội dung nghiên cứu

- Phân tích và hệ thống hóa các kỹ thuật Domain Generalization (DG).
- Xây dựng và triển khai mô hình MPCount gồm các thành phần
 - Cấu trúc Encoder-Decoder
 - Cơ chế Attention Memory Bank (AMB)
 - Content Error Mask (CEM)
 - Patch-wise Classification (PC)
- Đánh giá định lượng và phân tích hiệu năng
 - Thực hiện các thí nghiệm so sánh chéo giữa các tập dữ liệu
 - Nghiên cứu thành phần để làm rõ đóng góp của từng module

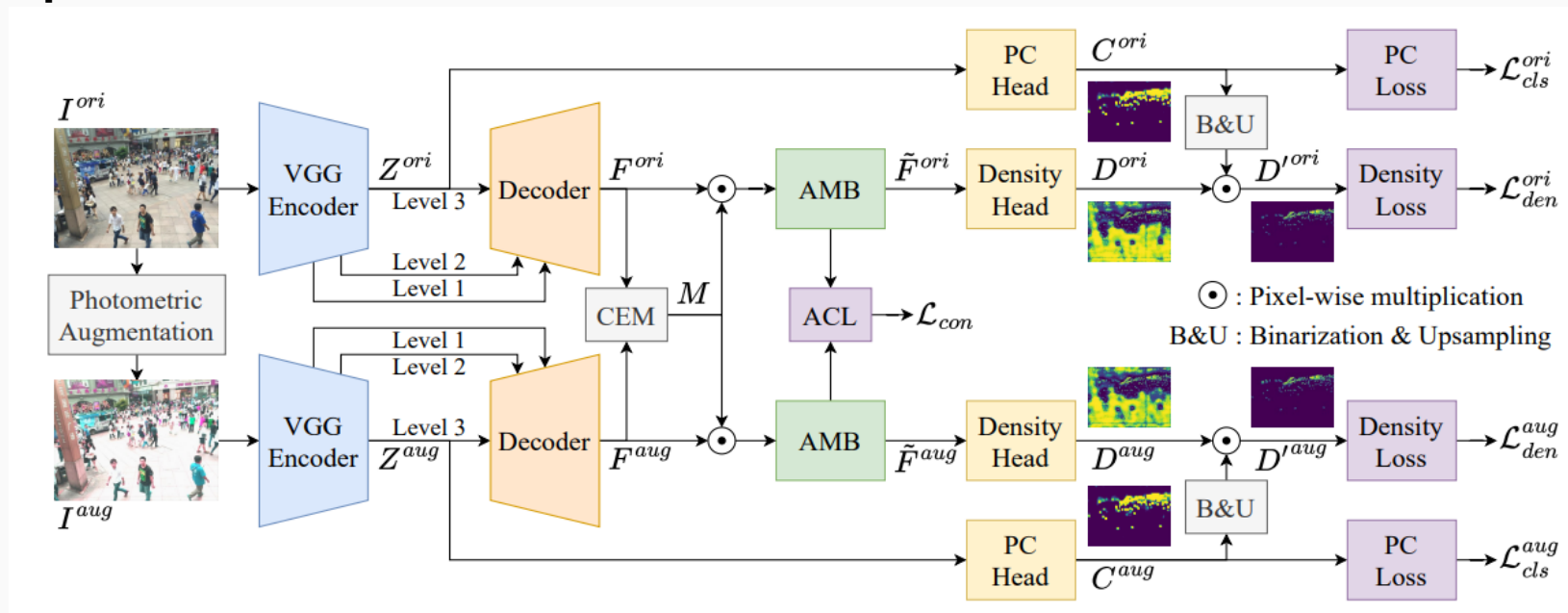
Nội dung và Phương pháp (2/3)

2. Phương pháp nghiên cứu

- **Phương pháp nghiên cứu tài liệu:** các công trình về Domain Generalization, các nghiên cứu crowd counting gần đây
- **Phương pháp thực nghiệm**
 - Triển khai mô hình MPCount bằng PyTorch
 - Huấn luyện trên các bộ dữ liệu công khai như ShanghaiTech Part A & B, JHU-Crowd++
- **Phương pháp đánh giá**
 - Chỉ số đánh giá: MAE, MSE
 - Thiết lập đánh giá: Huấn luyện trên một tập, kiểm tra cross-dataset
 - Phân tích bổ sung: Ablation Study - đánh giá vai trò của AMB, CEM, PC

Nội dung và Phương pháp (3/3)

Pipeline



Hình 3. Quy trình huấn luyện tổng thể của mô hình MPCount

Kết quả dự kiến

- **Về mặt lý thuyết:** Hệ thống hóa các kỹ thuật Domain Generalization trong bài toán đếm người
- **Về mặt mô hình và phương pháp:** Tái lập thành công khung mô hình MPCount
- **Về mặt thực nghiệm:** Đánh giá định lượng khả năng tổng quát hóa bằng MAE, MSE
- **Về mặt ứng dụng:** Nhận xét tiềm năng áp dụng trong các kịch bản thực tế

Tài liệu tham khảo

[1] Zhuoxuan Peng, S.-H. Gary Chan.

"Single Domain Generalization for Crowd Counting."

In proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024.

[2] Yingying Zhang, Desen Zhou, Siqin Chen, Shenghua Gao, and Yi Ma.

"Single-Image Crowd Counting via Multi-Column Convolutional Neural Network."

In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[3] Vishwanath A. Sindagi, Rajeev Yasarla, and Vishal M. Patel.

"JHU-CROWD++: A Large-Scale Benchmark for Crowd Counting."

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2020.

[4] Karen Simonyan and Andrew Zisserman.

"Very Deep Convolutional Networks for Large-Scale Image Recognition."

In International Conference on Learning Representations (ICLR), 2015.